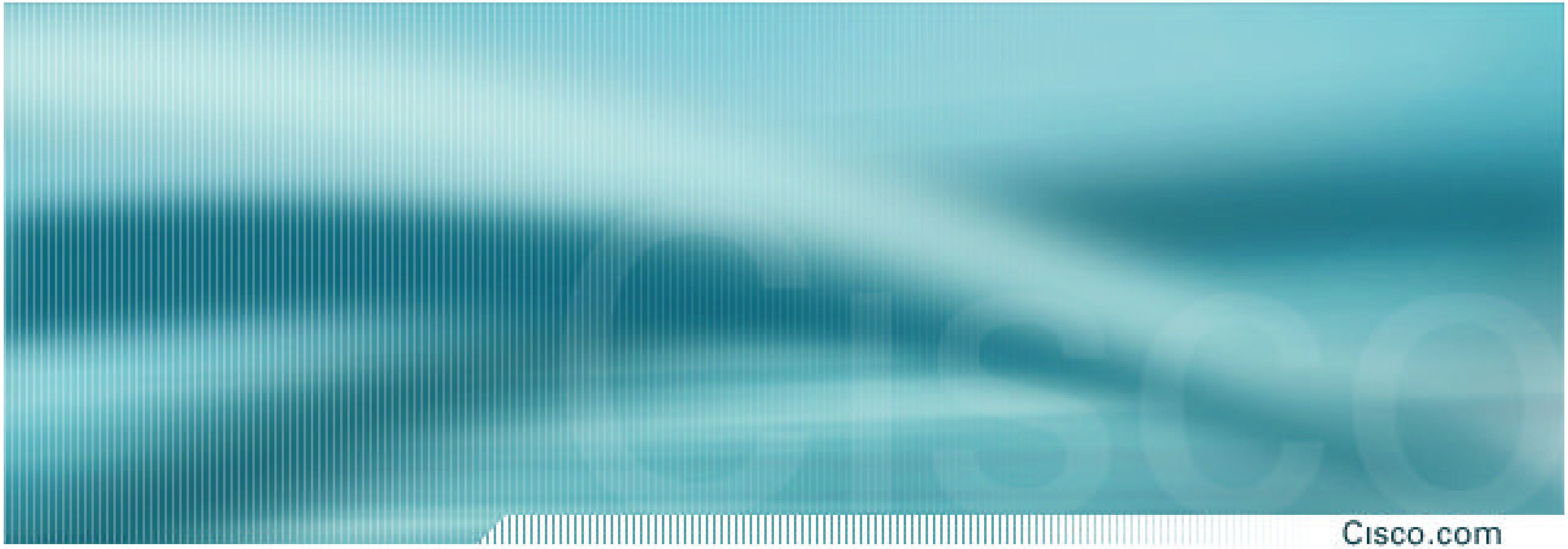
Cisco Systems ®



NetFlow Services

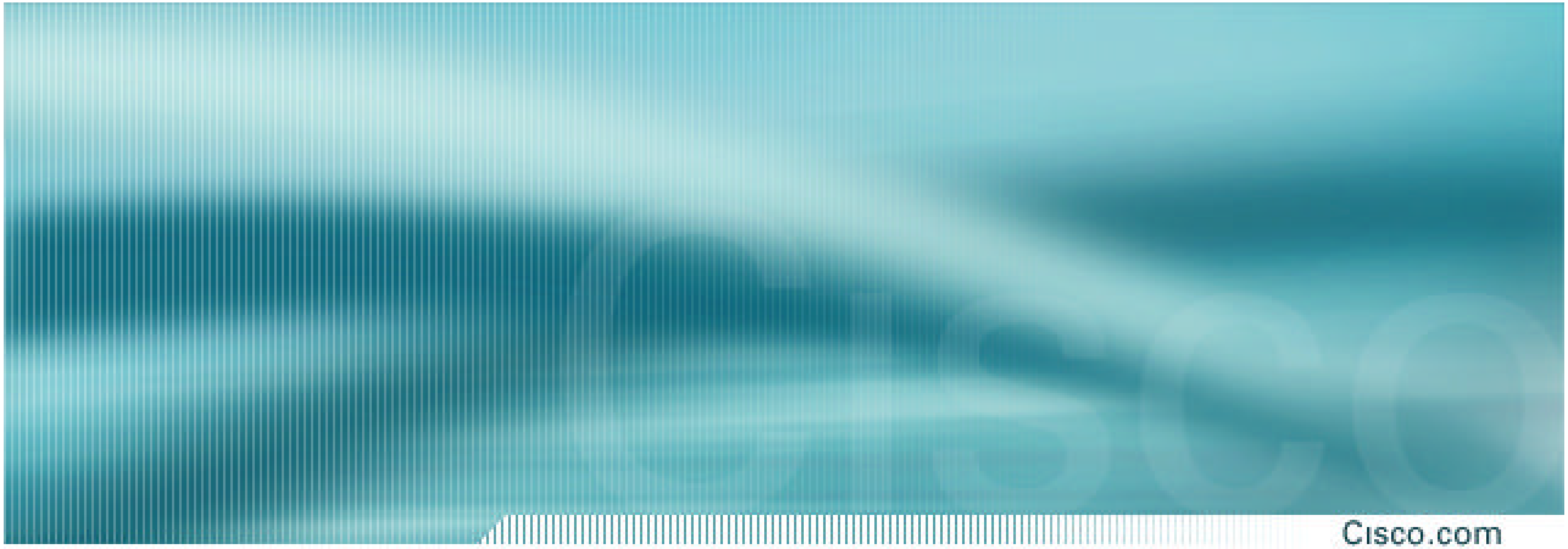
Benoit Claise

bclaise@cisco.com

RIPE 44, Amsterdam

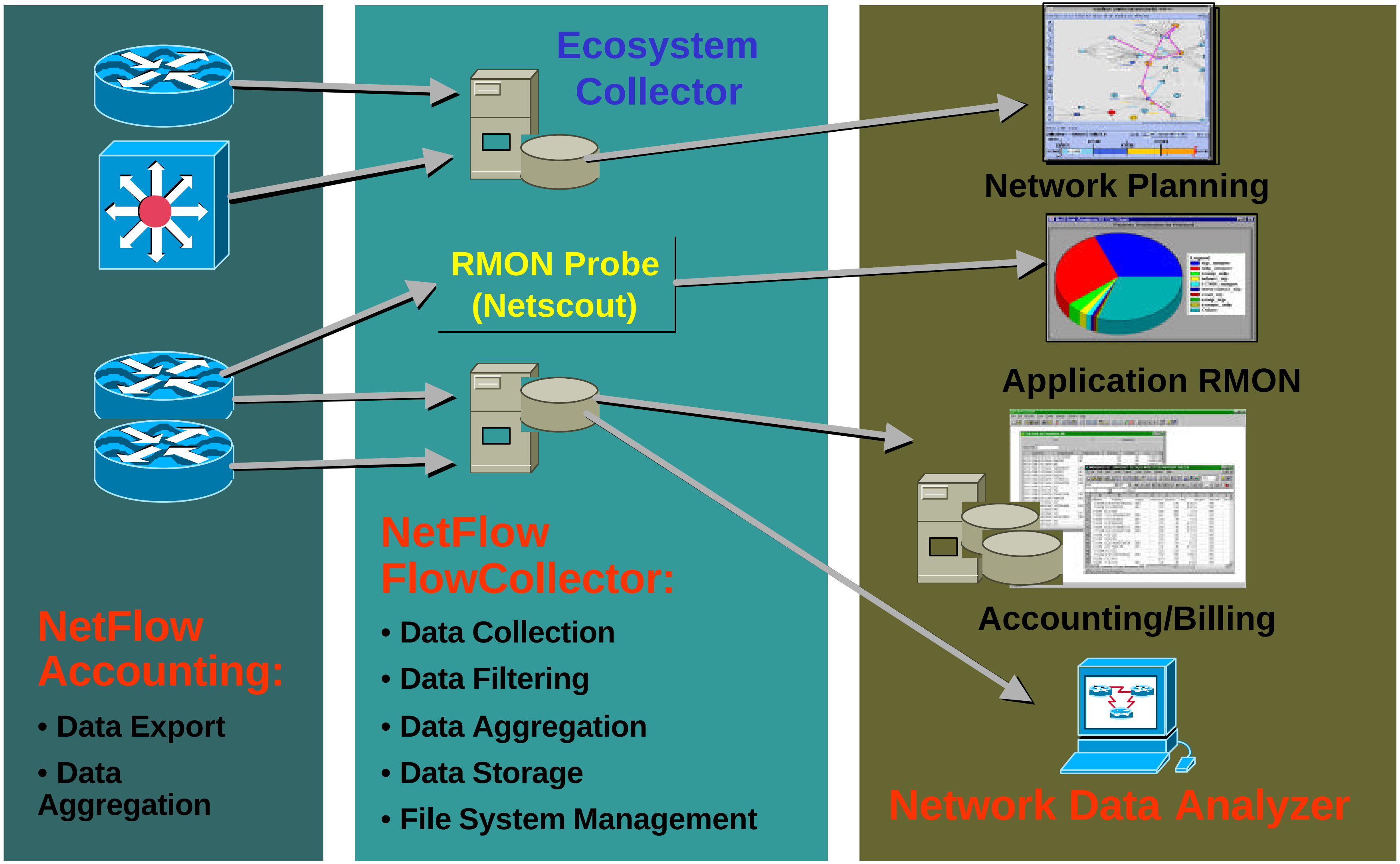
Table of Content

- NetFlow Basics
- NetFlow
    - Version 5 (Router)
    - Version 7 (Switch)
    - Version 8 (Router)
    - Sampled (12000 Series)
- Advanced Concepts
- Troubleshooting
- New Features
- New Features, Version 9 and the IETF
- New Features, MPLS Aware NetFlow
- New Features, BGP Next Hop Aggregation
- Roadmap
- NetFlow FlowCollector
- Deployment Guide



NetFlow Basics

NetFlow Infrastructure



NetFlow Accounting:

- Data Export
- Data Aggregation

NetFlow FlowCollector:

- Data Collection
- Data Filtering
- Data Aggregation
- Data Storage
- File System Management

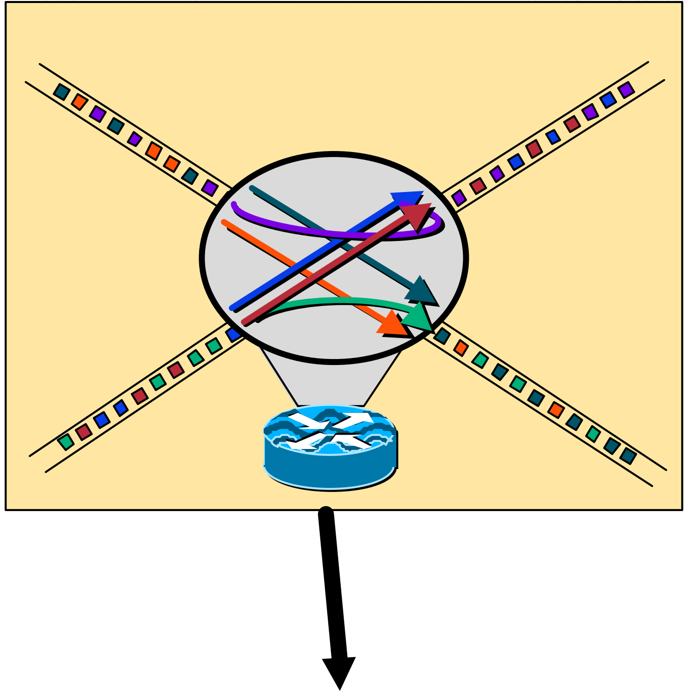
NetFlow Possible Applications

	NetFlow
Network Monitoring	X
Network Planning	X
Security Analysis	X
Application Monitoring	X
User Monitoring	X
Traffic Engineering	X
Peering Agreement	X
Usage-based Billing	X
Destination Sensitive Billing	X

What is a NetFlow Flow?

7 Keys define a flow

- Source Address
- Destination Address
- Source Port
- Destination Port
- Layer 3 Protocol Type
- TOS byte (DSCP)
- Input Logical Interface (ifIndex)



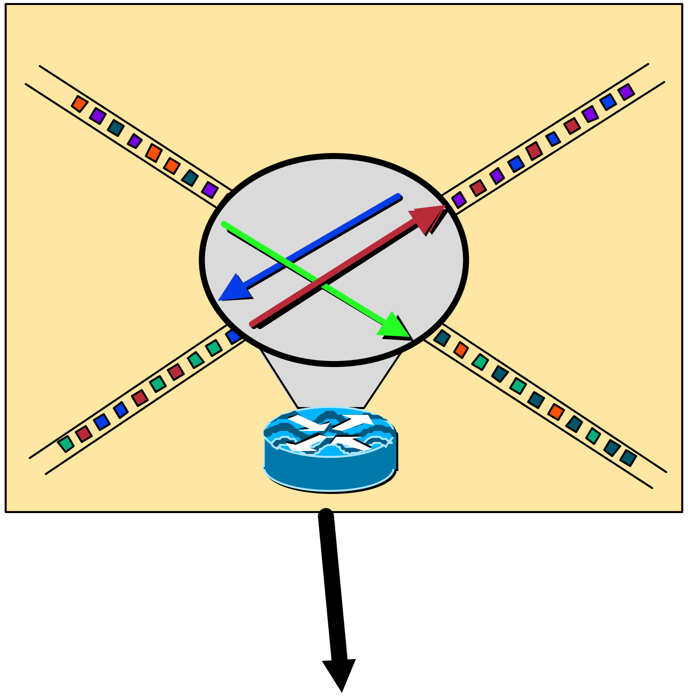
Exported Data

A flow is unidirectional

How does it work?

NetFlow Cache

7 identifiers	Other data
⋮	⋮

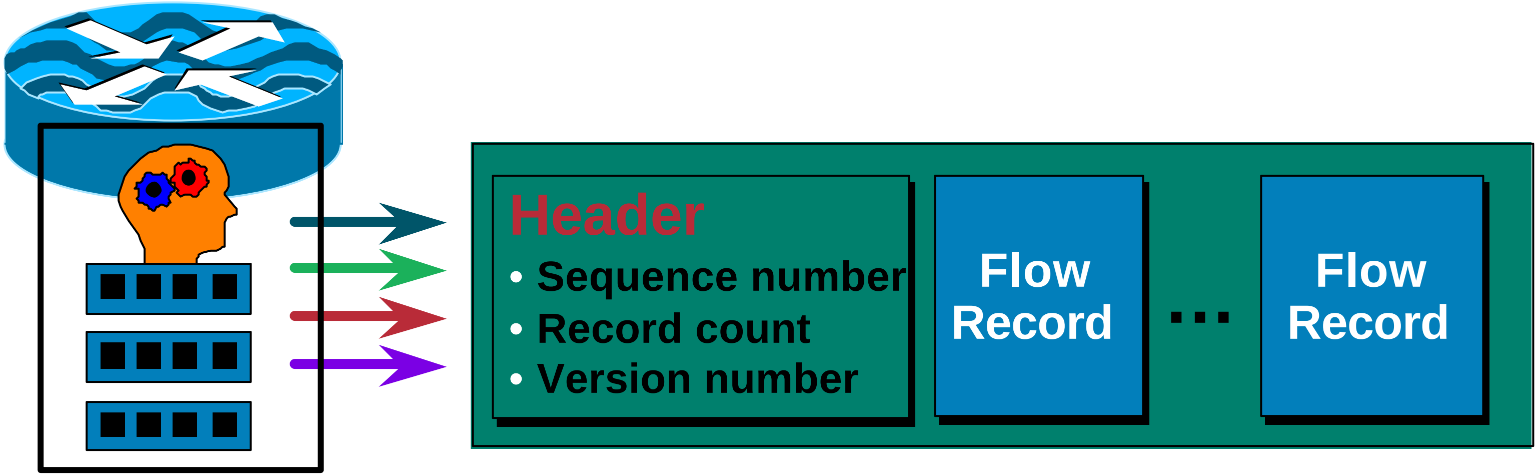


Exported Data

NetFlow Versions

- Version 5, the most complete version
- Version 7, on the switches
- Version 8, the Router Based Aggregation
- Version 9, the new flexible and extensible version

Data Export



Header
- Sequence number
- Record count
- Version number

Flow Record

…

Flow Record

NetFlow Cache

- Expired flows are grouped together into “Netflow Export” UDP datagrams for export to a collector
- UDP is used for speed and simplicity

NetFlow Principles

- Capture traffic statistics per port, protocol, BGP AS, network, …
- Support on most of the interface types
- Enable NetFlow on the main interface. But returns the sub-interface in the flow record (see new features)
- Supported on fast switching, Cisco Express Forwarding (CEF) and Distributed CEF

NetFlow Principles

- Not a switching path
- 7 flow identifiers
- Unidirectional traffic
- For ingress traffic only (*)
- IP unicast only (*)

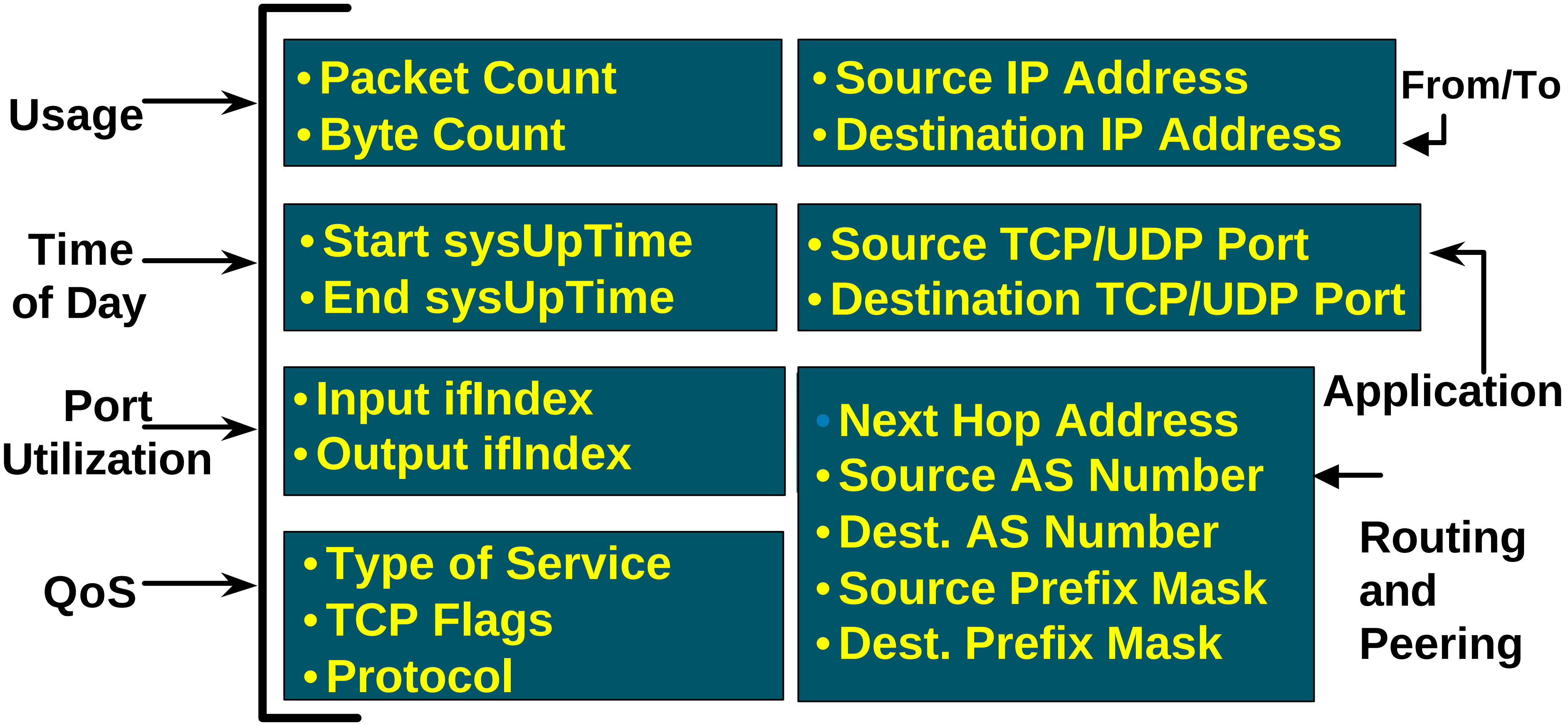
(*) See roadmap

NetFlow on the Router Version 5

Version 5

- Version 5 adds BGP AS
- Supported on router starting from 11.1 CA and 12.0
- The current version
- Note: No reason to use Netflow version 1 unless supporting a legacy collection system.

Version 5 Flow Format



Usage

- Packet Count
- Byte Count

- Source IP Address
- Destination IP Address

From/To

Time of Day

- Start sysUpTime
- End sysUpTime

- Source TCP/UDP Port
- Destination TCP/UDP Port

Application

Port Utilization

- Input ifIndex
- Output ifIndex

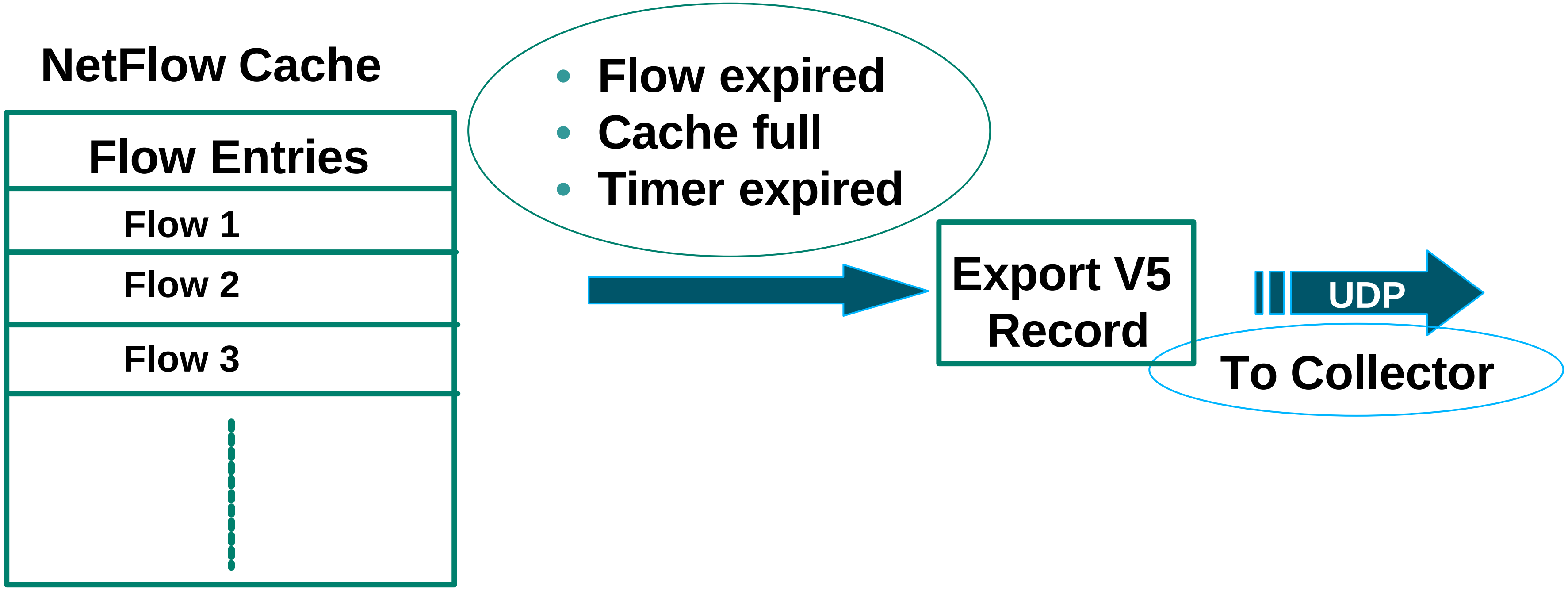
- Next Hop Address
- Source AS Number
- Dest. AS Number
- Source Prefix Mask
- Dest. Prefix Mask

Routing and Peering

QoS

- Type of Service
- TCP Flags
- Protocol

Version 5 Export



NetFlow Cache

Flow Entries

Flow 1

Flow 2

Flow 3

- Flow expired
- Cache full
- Timer expired

Export V5 Record

To Collector

Version 5 Configuration

```
router (config-if)#ip route-cache flow

router (config)#ip flow-export destination
  172.17.246.225 9996

router (config)#ip flow-export version 5 <peer-as |
  origin-as>
```

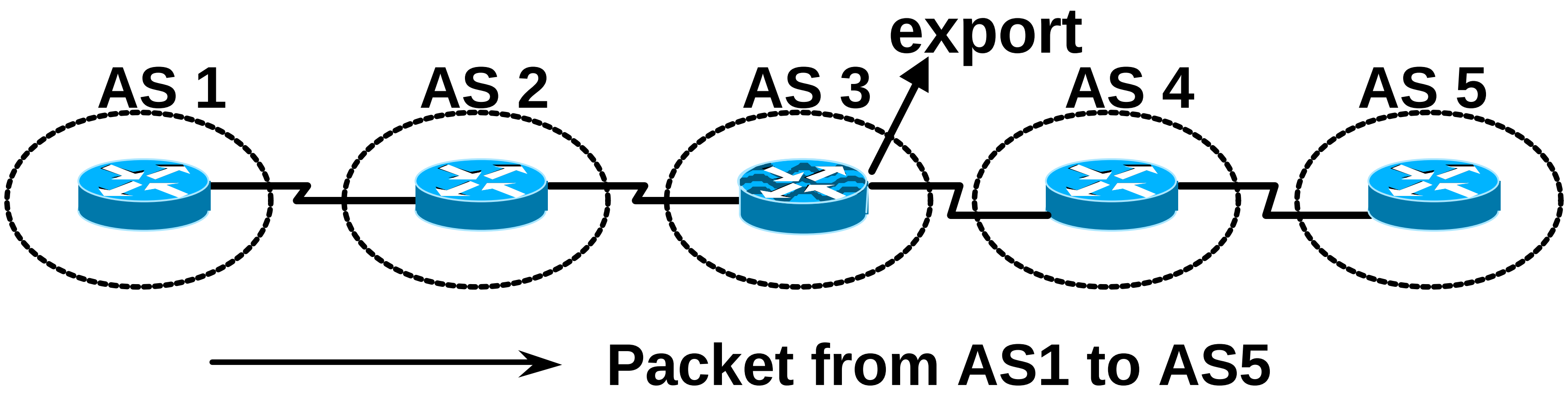
Optional configuration

```
router (config)#ip flow-export source loopback 0

router (config)#ip flow-cache entries <1024-524288>

router (config)#ip flow-cache timeout …
```

Version 5 Show Commands

```
martel#sh ip cache verbose flow
IP packet size distribution (94452 total packets):
   1-32   64   96  128  160  192  224  256  288  320  352  384  416  448  480
   .000 .199 .342 .300 .094 .028 .012 .005 .013 .000 .001 .000 .000 .000 .000

    512  544  576 1024 1536 2048 2560 3072 3584 4096 4608
   .000 .000 .000 .000 .000 .000 .000 .000 .000 .000 .000
IP Flow Switching Cache, 4456704 bytes
  1 active, 65535 inactive, 25322 added
  525430 ager polls, 0 flow alloc failures
  last clearing of statistics never
Protocol         Total    Flows   Packets Bytes  Packets Active(Sec) Idle(Sec)
--------         Flows     /Sec     /Flow  /Pkt     /Sec     /Flow     /Flow
TCP-BGP              7      0.0         2    41      0.0       1.6       7.5
UDP-TFTP             1      0.0         1    67      0.0       0.0      15.1
UDP-other        19884      0.0         3   111      0.1       5.6      15.4
ICMP              5429      0.0         3    41      0.0       0.9      15.5
Total:           25321      0.0         3    97      0.2       4.6      15.4

SrcIf          SrcIPaddress    DstIf          DstIPaddress    Pr TOS Flgs  Pkts
Port Msk AS                    Port Msk AS    NextHop              B/Pk  Active
Se0/1          193.1.1.3       Se0/0          172.17.246.228  11 00   10      5
00A1 /24 193                   C628 /0   0    0.0.0.0                 84   39.7
```

Origin Autonomous System



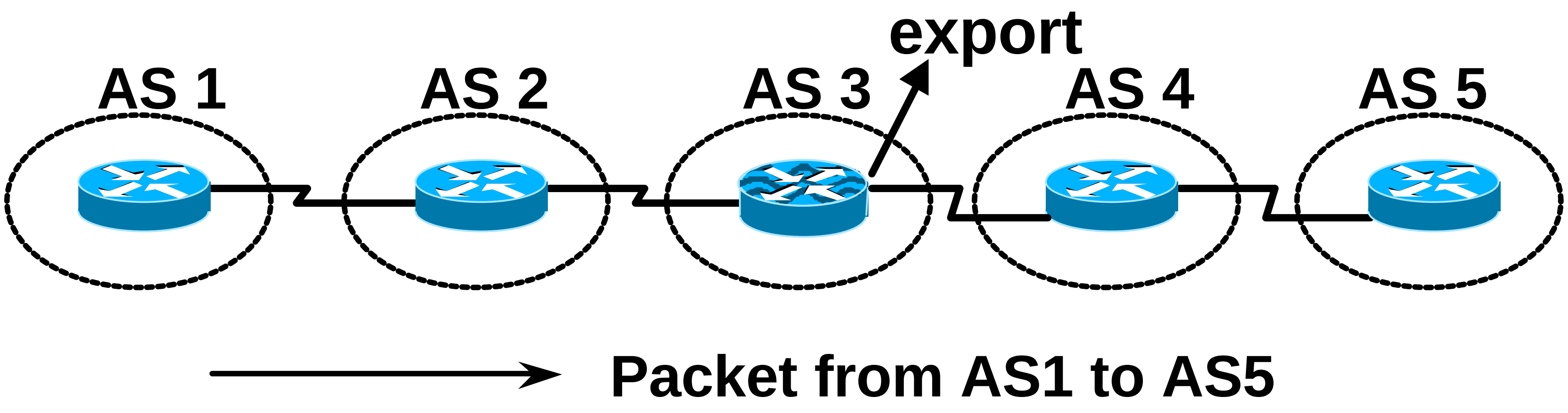
- **ip flow-export version 5 origin-as**

  Source AS: AS1

  Destination AS: AS5

- **Important: the AS fields will stay empty with only "ip flow-export version 5"**

Peer Autonomous System



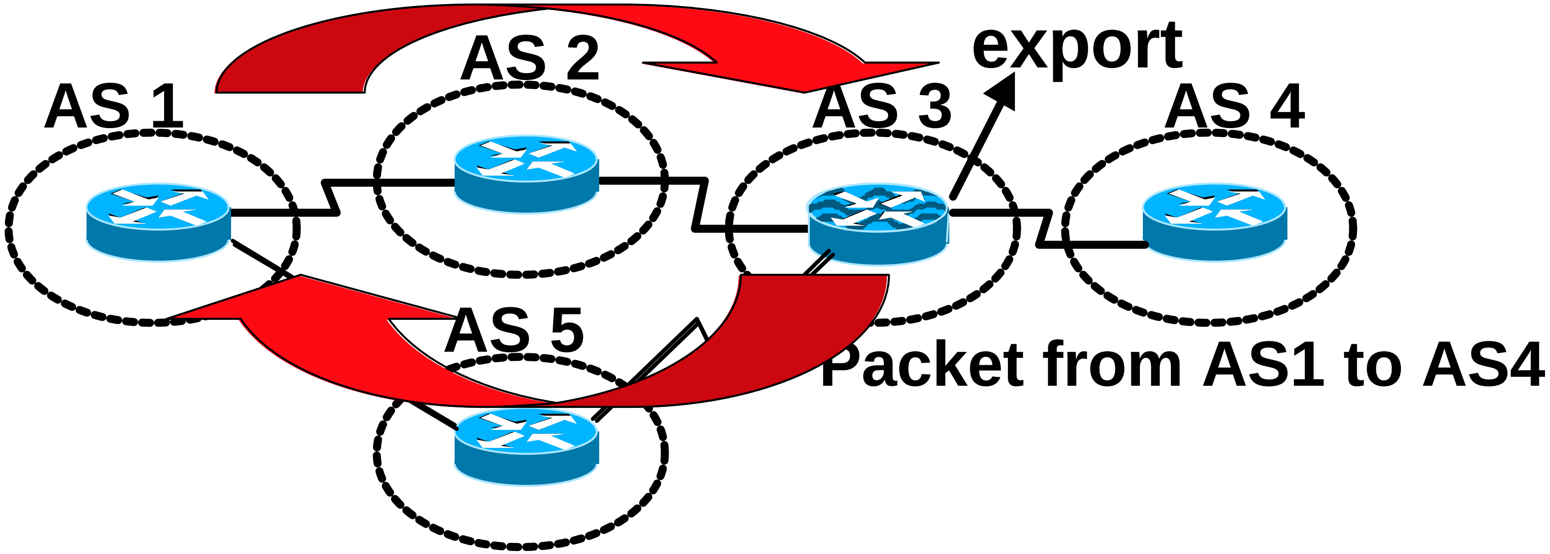
- **ip flow-export version 5 peer-as**

  Source AS: AS2

  Destination AS: AS4

- **Important: the AS fields will stay empty with only "ip flow-export version 5"**

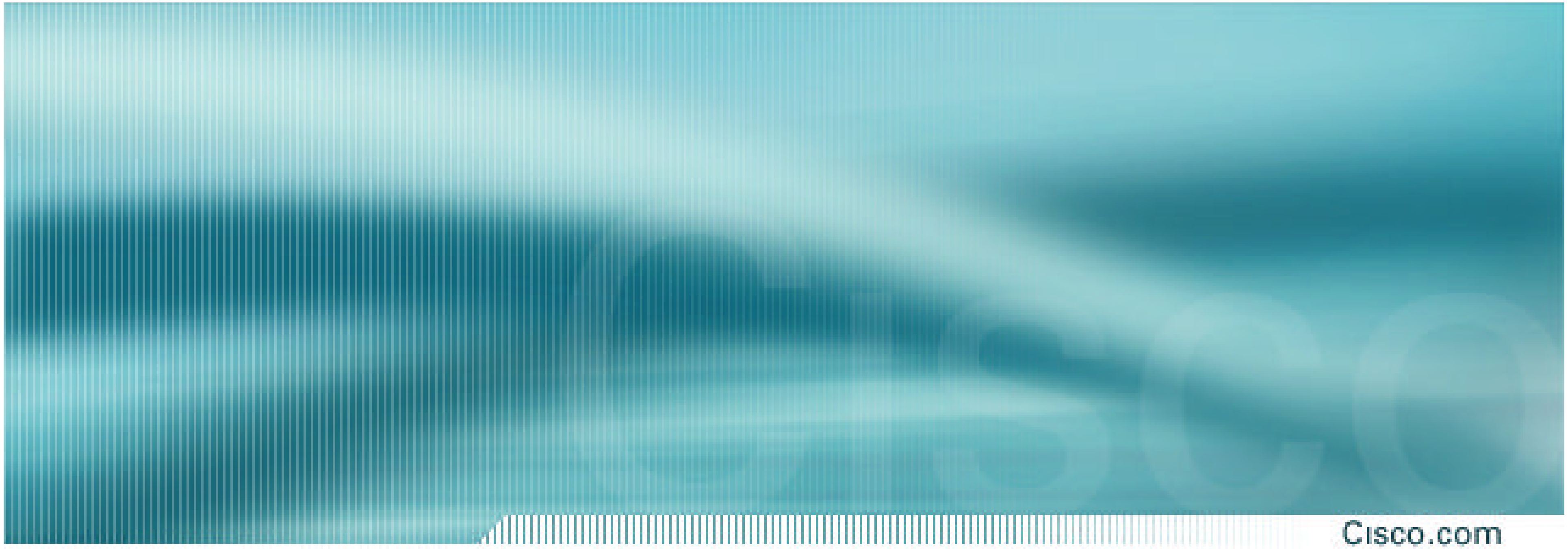
Asymetric BGP traffic Problem



Origin-as: AS1 and AS4 CORRECT

Peer-as: AS5 and AS4 WRONG

Because of the source IP address lookup in the BGP table

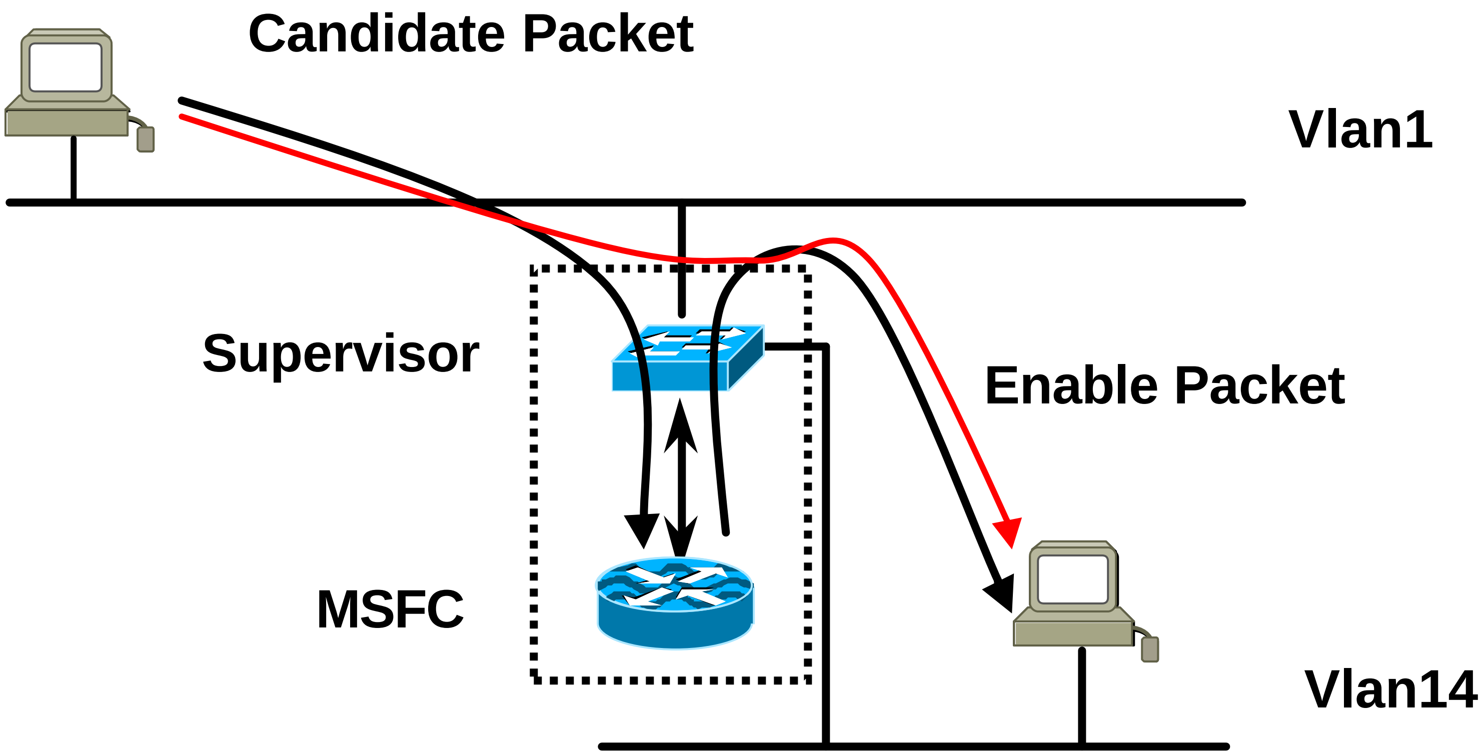


NetFlow on the Switches Version 7

NetFlow Version 7

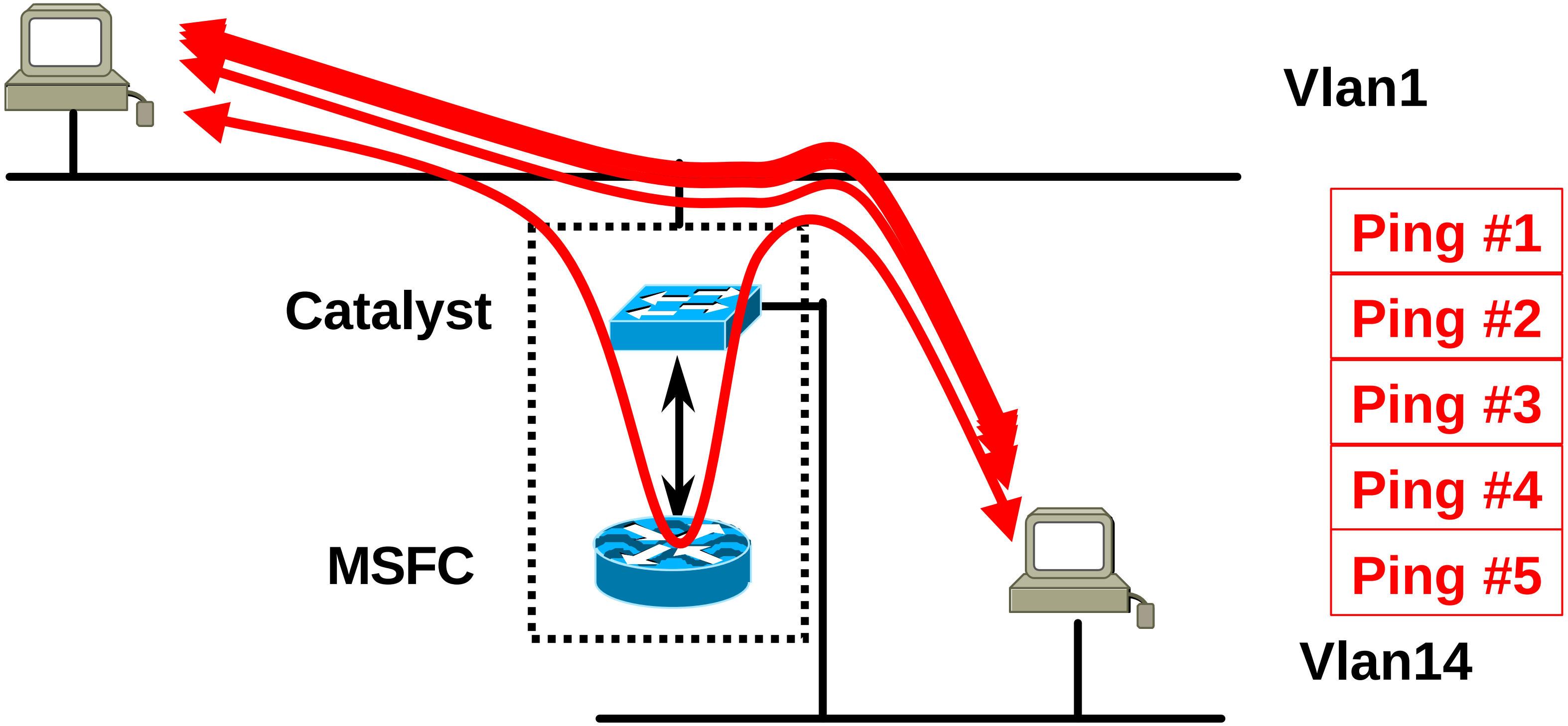
- Support for Catalyst switches with a layer 3 board:
    - Catalyst 5000 with a RSM (Route Switch Module)
    - Catalyst 6000 with a MSFC (MultiLayer Switching Feature Card)
- Version 7 uses MultiLayer Switching (MLS) or CEF with a catalyst 6000 with SUP2
- For IP unicast only, not multicast, not IPX, even if MLS can do all three
- MLS cache equals to the NetFlow cache. Confusion in the documentation

MLS Example



Layer 3 Switched

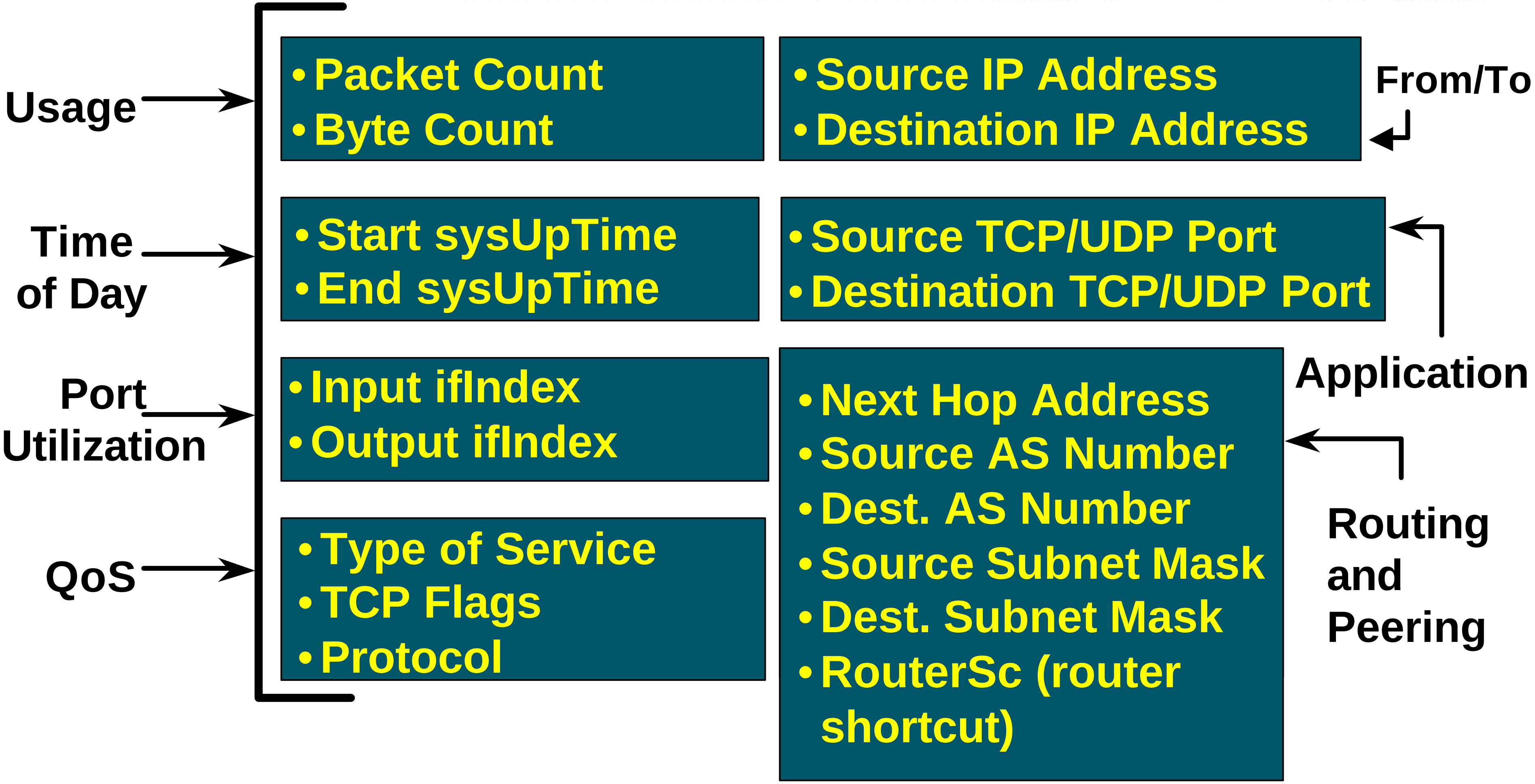
MLS Example



MLS Concepts

- MLS is enabled for the whole device, not per interface like on a router. So no concept of incoming/outgoing traffic
- MLS is not for layer 2 traffic (see new features)
- MLS export the layer 3 information
- The MLS switching is done in hardware for the catalyst (5000/6000). Which means that only the export takes some CPU

Version 7 Flow Format



Usage

- Packet Count
- Byte Count

Time of Day

- Start sysUpTime
- End sysUpTime

Port Utilization

- Input ifIndex
- Output ifIndex

QoS

- Type of Service
- TCP Flags
- Protocol

From/To

- Source IP Address
- Destination IP Address

Application

- Source TCP/UDP Port
- Destination TCP/UDP Port

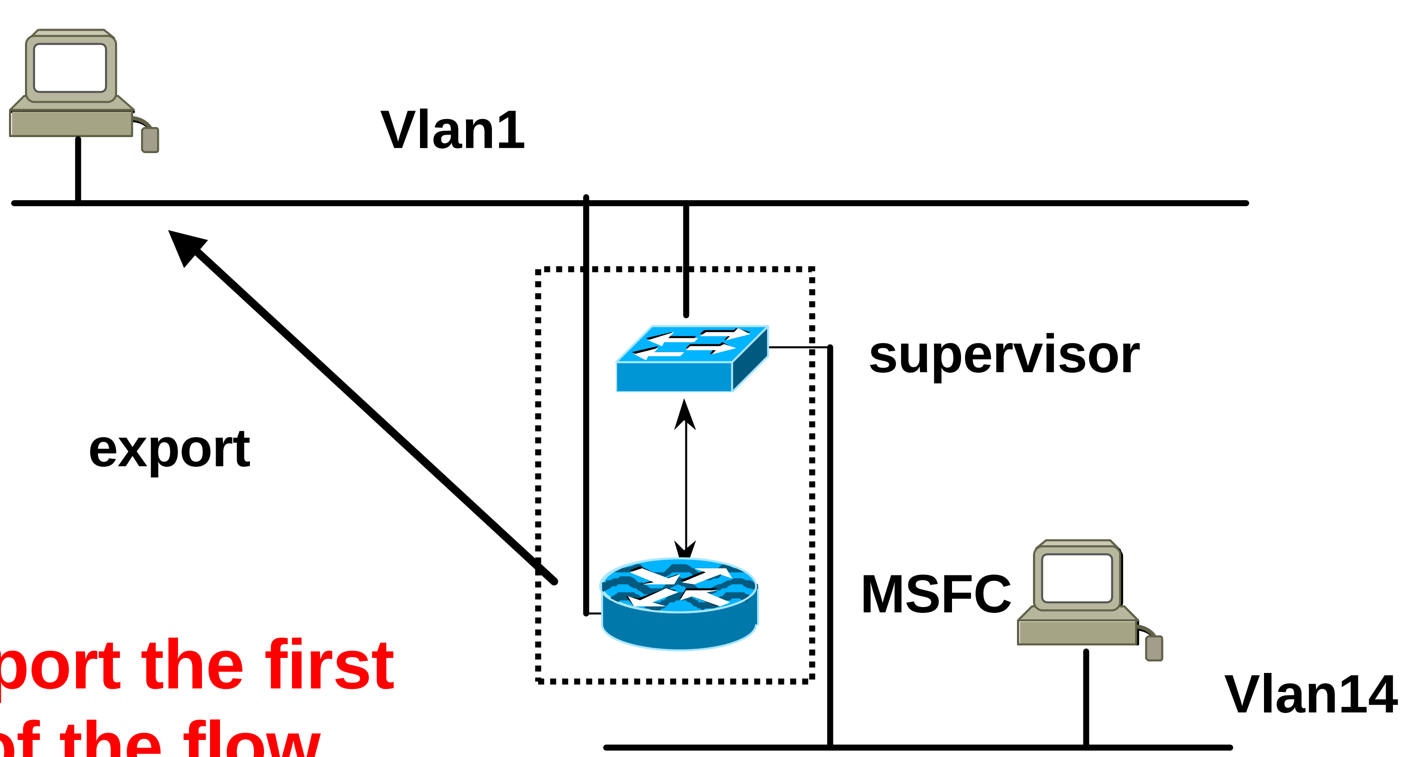
Routing and Peering

- Next Hop Address
- Source AS Number
- Dest. AS Number
- Source Subnet Mask
- Dest. Subnet Mask
- RouterSc (router shortcut)

Note that some of fields are not populated

Bad Design

MLS/NDE (not) enabled and export v5 from the MSFC

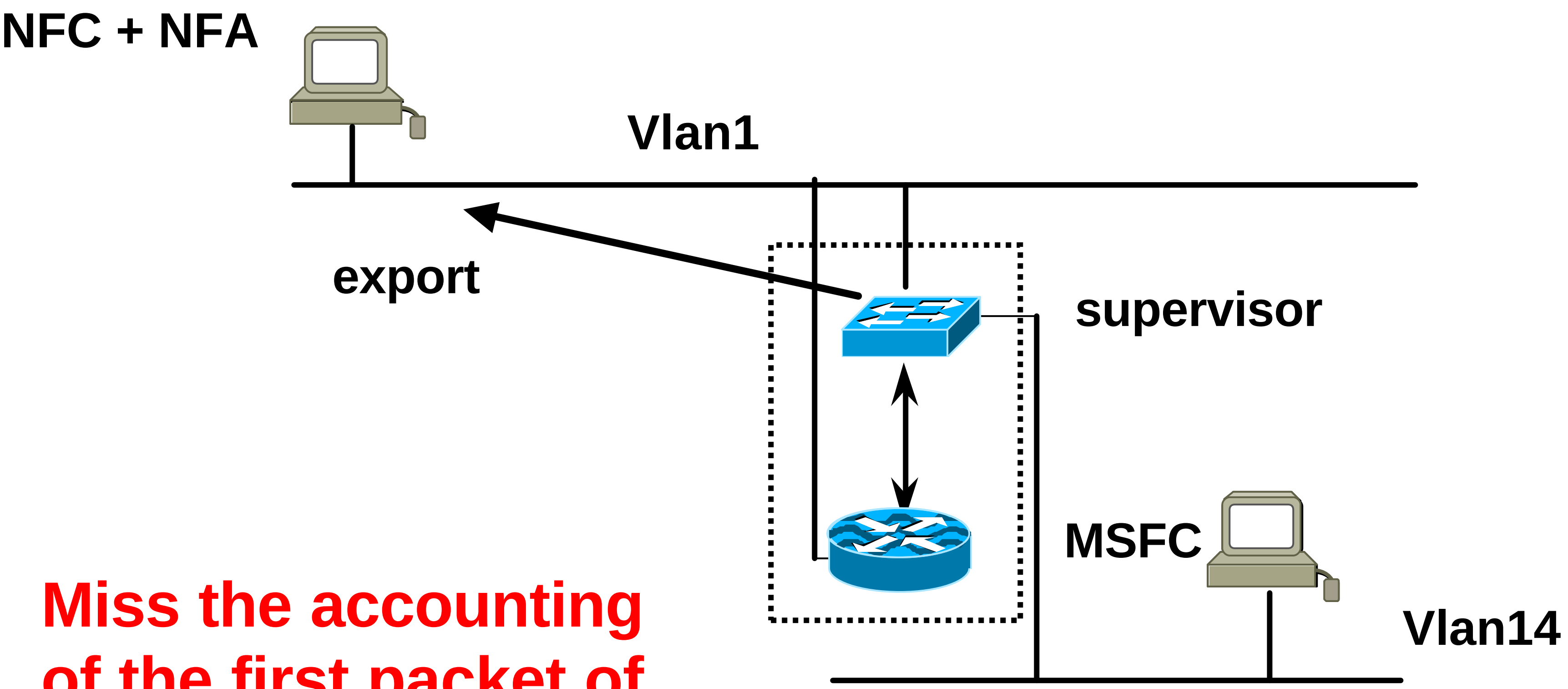


Only export the first packet of the flow

Unless don’t use MLS…

Approximate Design

MLS/NDE enabled and export v7 from the SUP

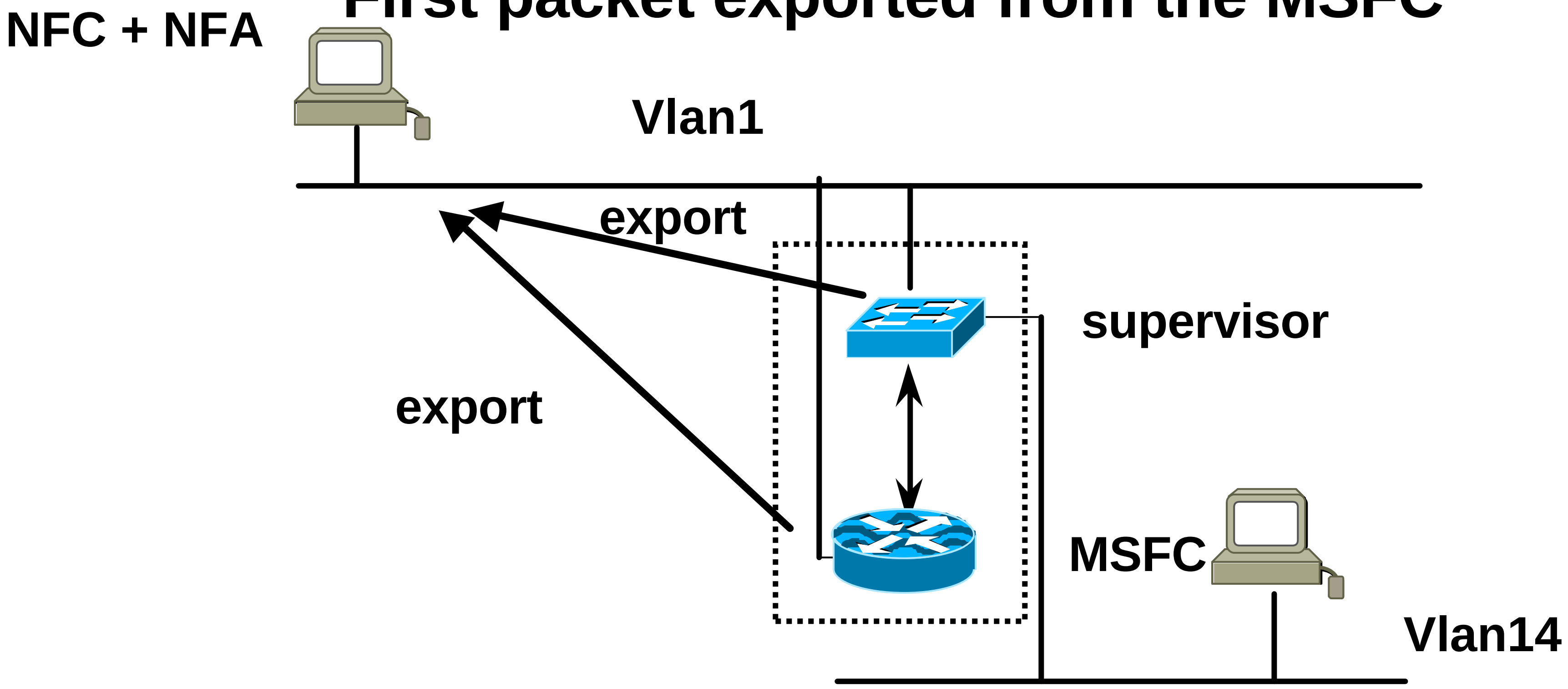


Miss the accounting of the first packet of the flow

Better Design

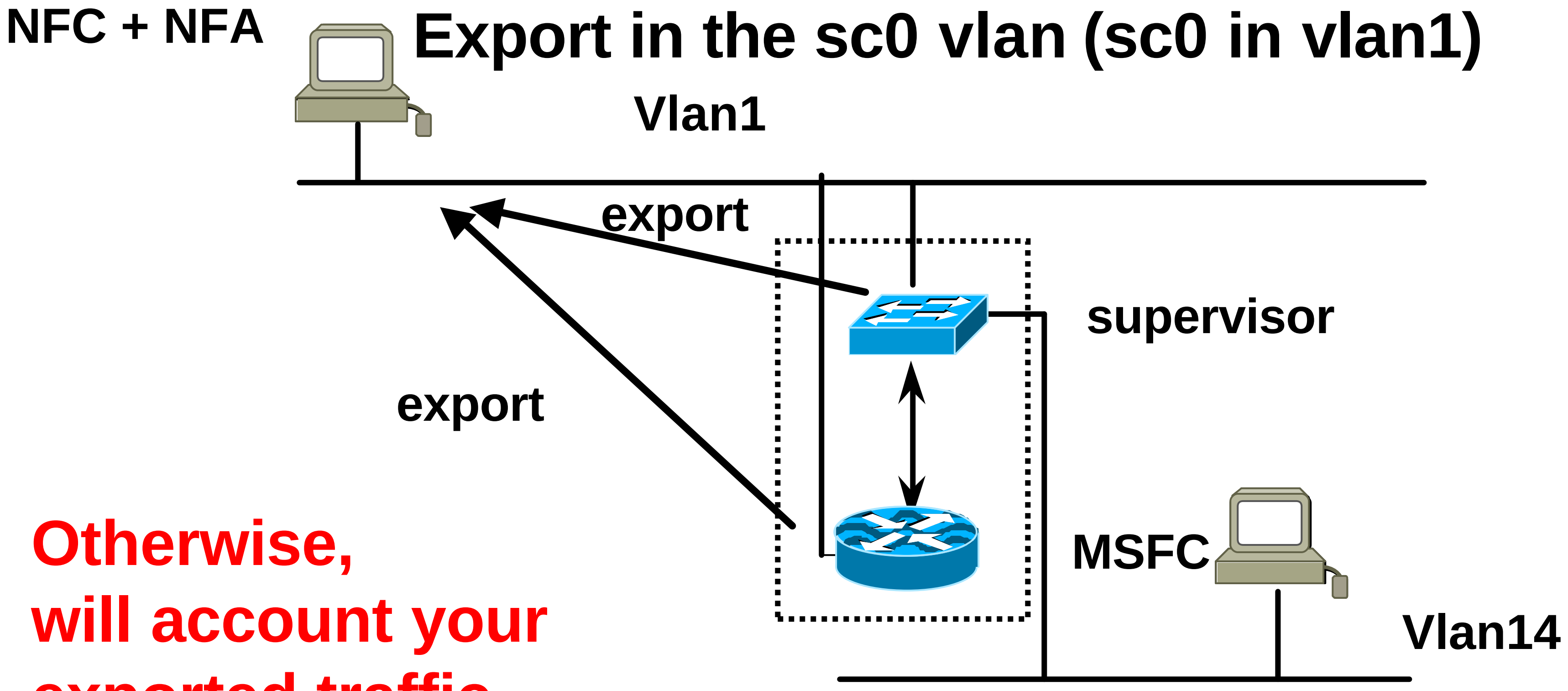
MLS/NDE enabled and export v7 from SUP
export v5 from the MSFC

First packet exported from the MSFC



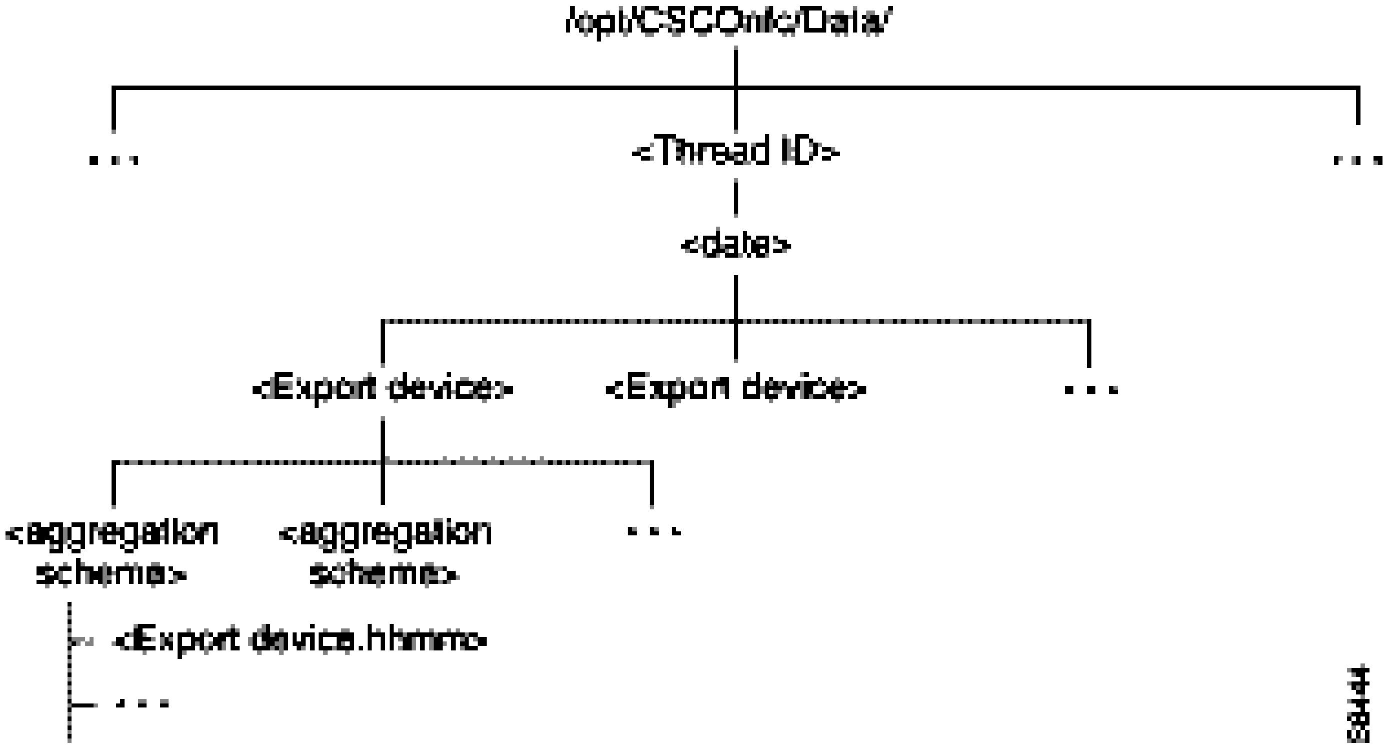
Best Design

MLS/NDE enabled and export v7 from the SUP
export v5 from the MSFC
First packet exported from the MSFC
Export in the sc0 vlan (sc0 in vlan1)



Otherwise,
will account your
exported traffic

Best Design Problem



- The Collector doesn’t correlate the flows from the same physical device
- The 2 different directories will be created

Best Design Solution

```
# In case of V7, set
  USE_SHORT_CUT_ADDRESS_AS_SOURCE_IP to
  "yes" so that FlowCollector will use the
  address of the router being short-cut as
  the source of the corresponding flow.
  Default is set to No

USE_SHORT_CUT_ADDRESS_AS_SOURCE_IP No
```

- **Change the nf.resources configuration file**

The Cat6000

- Hybrid mode (catOS/IOS) or native mode (full IOS)
- MLS is internal (no external MLS RP)
- SUP1 or SUP2, MSFC1 or MSFC2, PFC1 or PFC2
- In PFC1, uses MLS: a cache-based scheme
- In PFC2, uses HW CEF implementation, with a FIB: PFC2 comes with MSFC2 and SUP2

Cat6000 with a SUP2

- The PFC2 (on the SUP2) uses CEF, not MLS anymore
- We still have the NetFlow for accounting only, next to the Forwarding Information Base
- Cisco Express Forwarding (CEF) overview

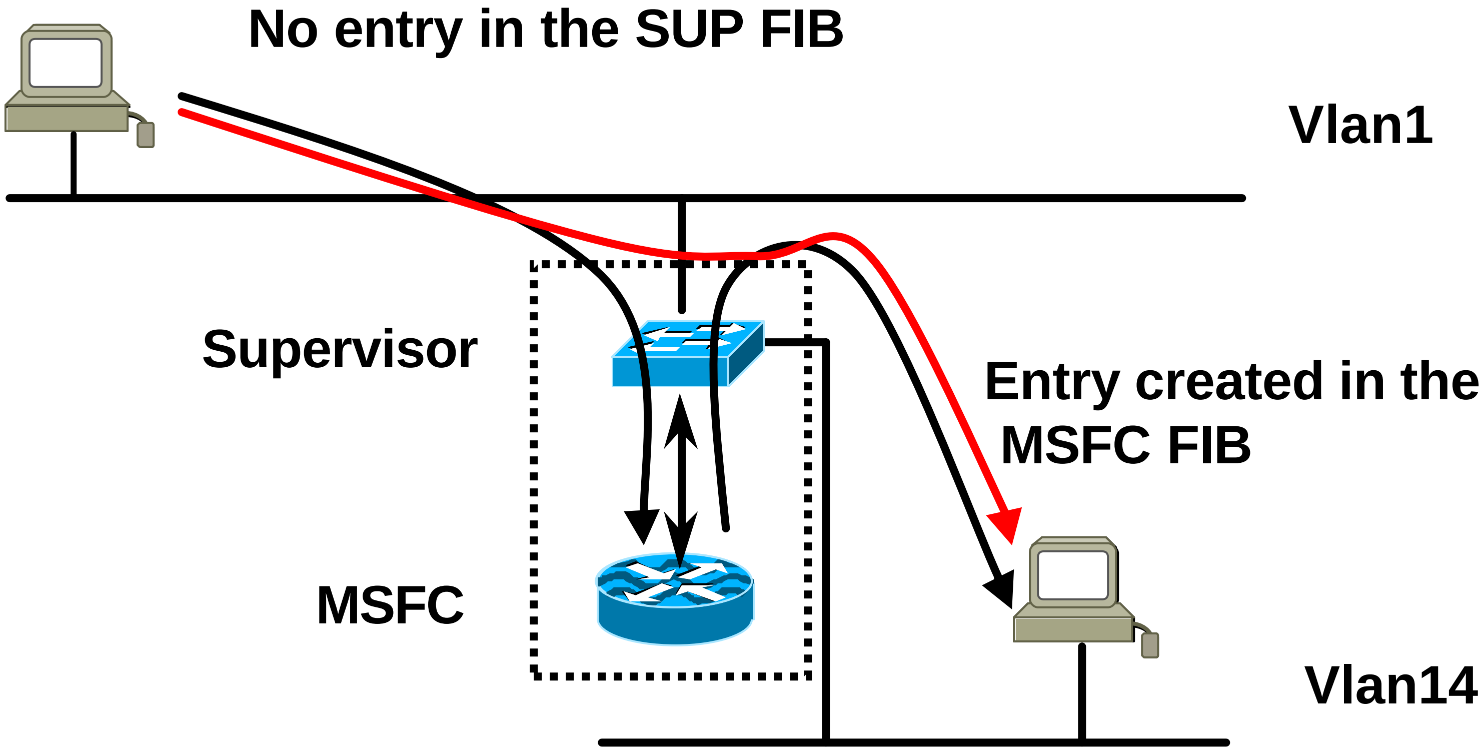
  CEF: No route cache, the router maintains a Forwarding Information Base (FIB) which is a mirror of the routing table

  Uses Forwarding Information Base (FIB) for route lookup and adjacency for encapsulation

  FIB synchronisation between the MSFC and the supervisor

DCEF Example

FIB Synchronisation



All entries go through the SUP FIB

Cat6000 with a SUP2, CEF mechanism

- **Test of 5 inter vlans pings through a cat6000**
- **The dest. host has no adjacency in the FIB**
- **The first packet is sent to the MSFC for the ARP request to be sent in the correct vlan.**
  This packet is not accounted by the SUP
- **If NetFlow is enabled on the MSFC, this packet will be accounted**
- **ARP reply arrives and updates MSFC FIB**
- **The MSFC FIB updates the SUP FIB**
- **The 4 next pings go through and are accounted by the SUP version 7 export**

Cat6000 with a SUP2, Export or Not on the MSFC?

- (-) Will account ONLY the first packet of a destination, the one which will complete the glean adjacency
- (-) The FIB entries remain the time of the ARP entries. Not updated so often as the MLS entries!

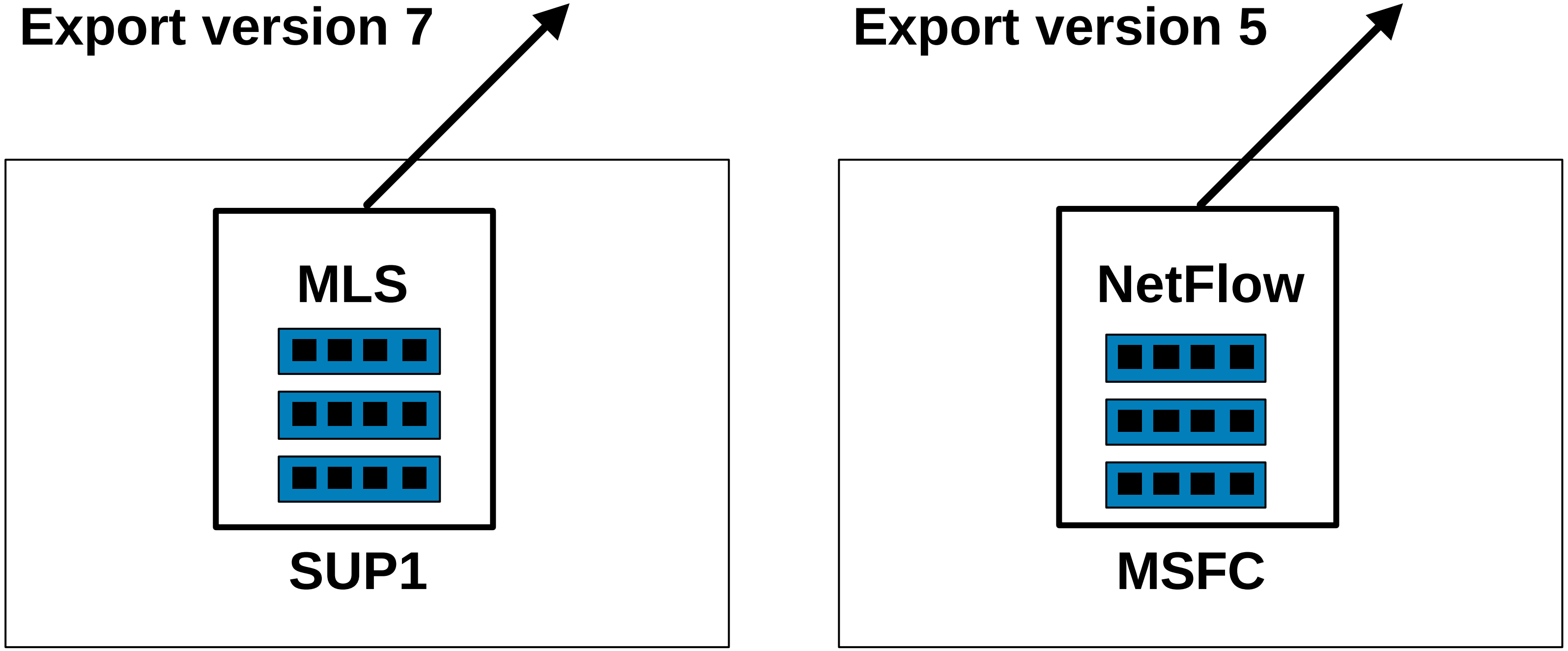
Cat6000 with a SUP2, Export or Not on the MSFC?

- (+) Will account the first packet of a destination, the one which will complete the glean adjacency
- (+) Some features still use MLS
- (+) Some features will always go through the MSFC: NAT, IP access-list with log, etc…
- Conclusion:

  The export is needed for accounting accuracy

  But less important as for MLS with a SUP1

Caches – Cat6000



Export version 7

MLS

SUP1

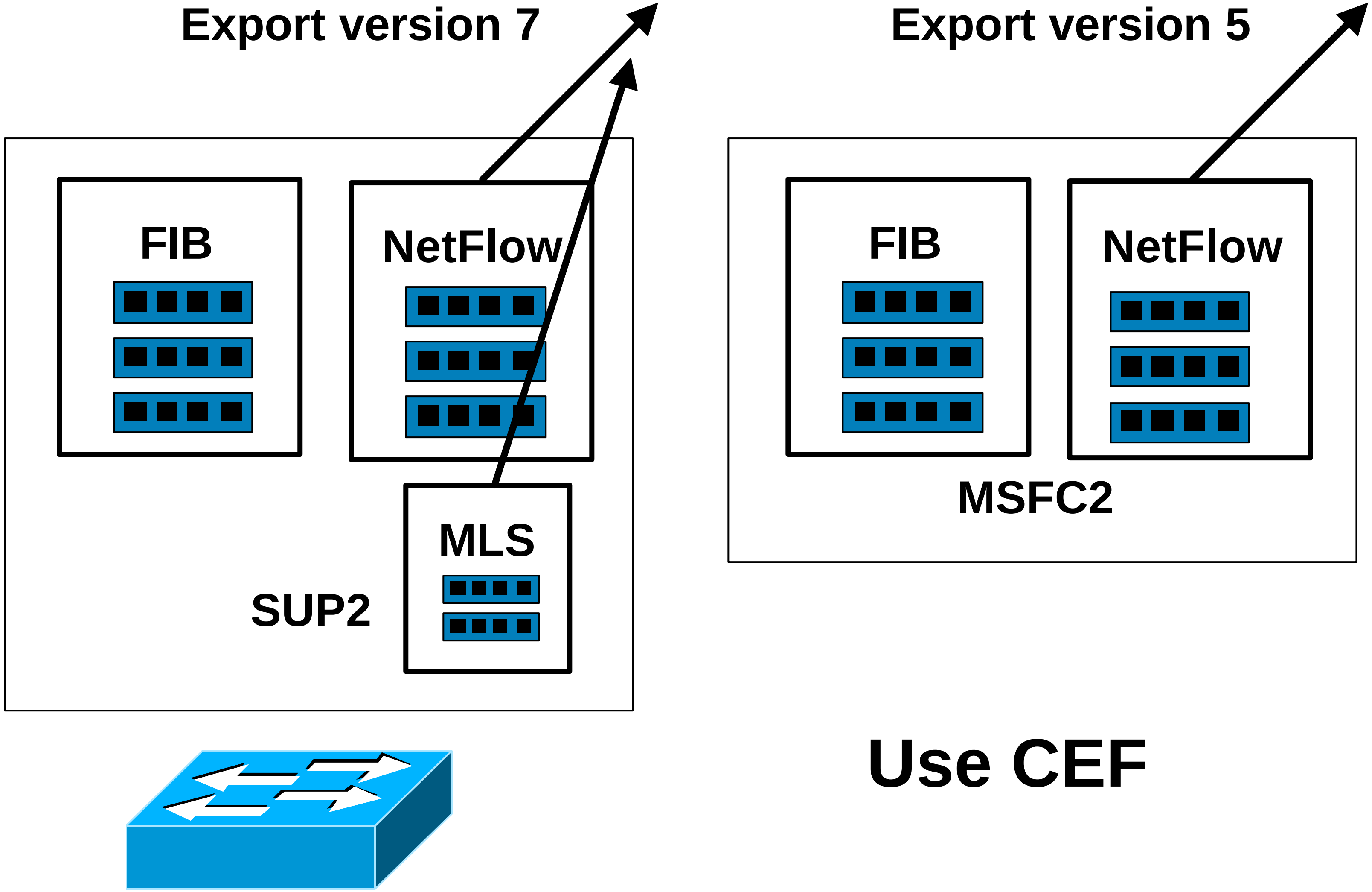
Export version 5

NetFlow

MSFC

Use MLS

Caches – Cat6000 with SUP2/PFC2



Use CEF

Cat6000, Native Mode

```
mls flow ip full                                     -> flow mask
mls nde src_address 10.200.8.127 version 7
   -> version 7 export source OR
mls nde sender   -> NDE enable + NDE from the PFC uses the
 source configured from the MSFC!!!!!
interface vlan 1
  ip address 10.100.8.127 255.255.255.0
  ip route-cache flow
interface FastEthernet 3/2
  ip address 10.200.8.2 255.255.255.0
  ip route-cache flow

ip flow-export source vlan1    -> version 5 export source
ip flow-export version 5
ip flow-export destination 172.17.246.244 9996
                         -> both for version 5 and 7 export
```

Cat6000, Native Mode

```
Cosmos#sh mls nde
Netflow Data Export enabled
Netflow Data Export configured for port 9996 on Host
  172.17.246.244
Source address: 10.200.8.127, port: 50191
Version: 7
  Include Filter not configured
  Exclude Filter not configured
  Total Netflow Data Export Packets are:
    3 packets, 0 no packets, 23 records
```

Cat6000, Native Mode

```
Cosmos#sh ip flow-export
exportFlow export is enabled
Exporting flows to 172.17.246.244 (9996)
Exporting using source interface Vlan1
Version 5 flow records
317 flows exported in 218 udp datagrams
0 flows failed due to lack of export packet
60 export packets were sent up to process level
0 export packets were dropped due to no fib
0 export packets were dropped due to adjacency issues
```

Format Comparison

Content	V5	V7
Source IP address	•	zero in case of destination-only
Destination IP address	•	•
Source TCP/UDP Port	•	zero in case of destination-only or source-destination
Destination TCP/UDP Port	•	zero in case of destination-only or source-destination
Next Hop Router IP address	•	always zero
Input Physical Interface Index	•	It depends
Output Physical Interface Index	•	It depends
Packet Count for this flow	•	•
Start of Flow Timestamps	•	•
End of Flow Timestamps	•	•

New

New

New

Format Comparison

Content	V5	V7
IP Protocol (TCP=6, UDP=17)	·	zero in case of destination-only or source-destination
Type Of Service byte	·	switch sets it to the TOS of first packet in flow
TCP flags	·	always zero
Source AS number	·	always sero
Destination AS number	·	always zero
Source Subnet Mask	·	always zero
Destination Subnet Mask	·	always zero
Flags (indicate invalid field within the flow)		·
Shortcut Router IP address		·

New

New

New Features

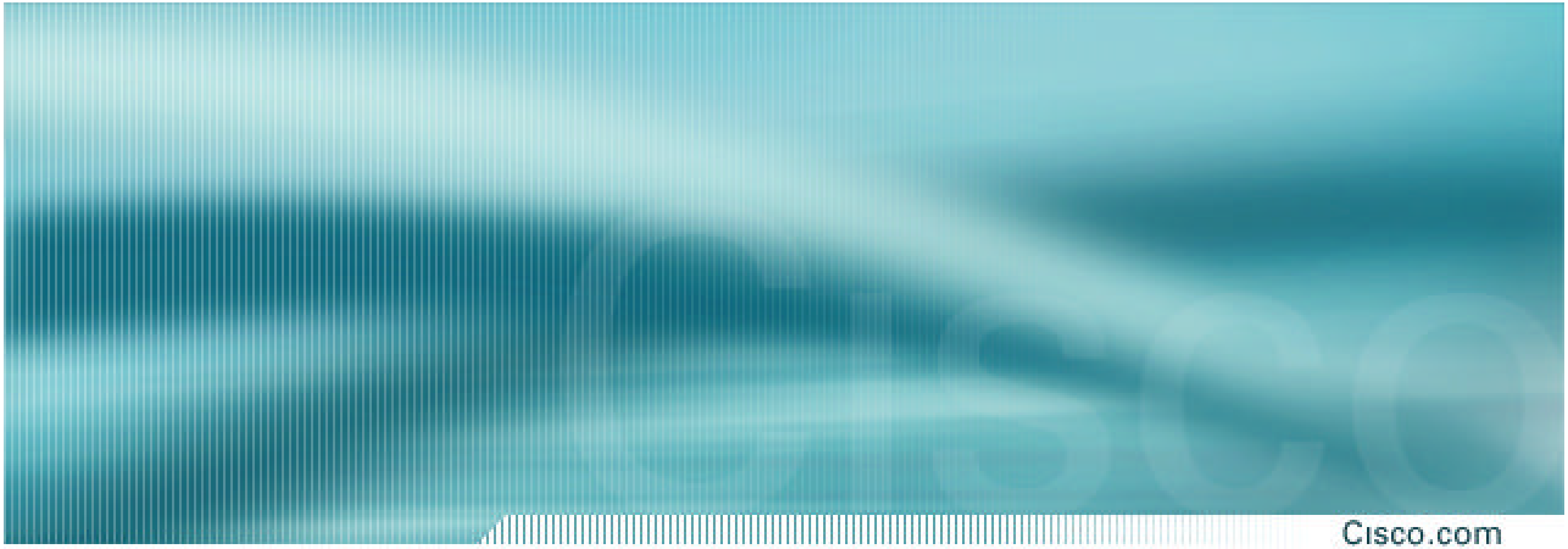
- **SUP2/PFC2 (EARL6) supports from 12.1(13)E:**

  Source and Destination BGP AS

  Input and Output ifIndexes

  Next Hop

- **Note: 12.1(13)E1 if any WAN cards**



NetFlow on the Router Version 8

Introduction

- Router Based Aggregation, i.e. version 8
- Enables router to summarize NetFlow data
- Reduces NetFlow Export data volume
- Decreases NetFlow Export bandwidth requirements
- Making collection easier

Introduction

- Supported from 12.0(3)T, 12.0(3)S and 12.1 On-board aggregation, the router maintains extra NetFlow cache(s), just for accounting.
- Still needs the main cache (version 5)
- When flows expire from the main cache, they are added to each enabled aggregation cache
- Several aggregations can be enabled at the same time

Aggregations

- Currently 5 aggregations: ProtocolPort, AS, SourcePrefix, DestinationPrefix, Prefix
- 6 extra aggregations available in IOS 12.0(15)S, Targeted for 12.2(1)T, containing the TOS
- Requires the new NetFlow Collector 3.5 or above

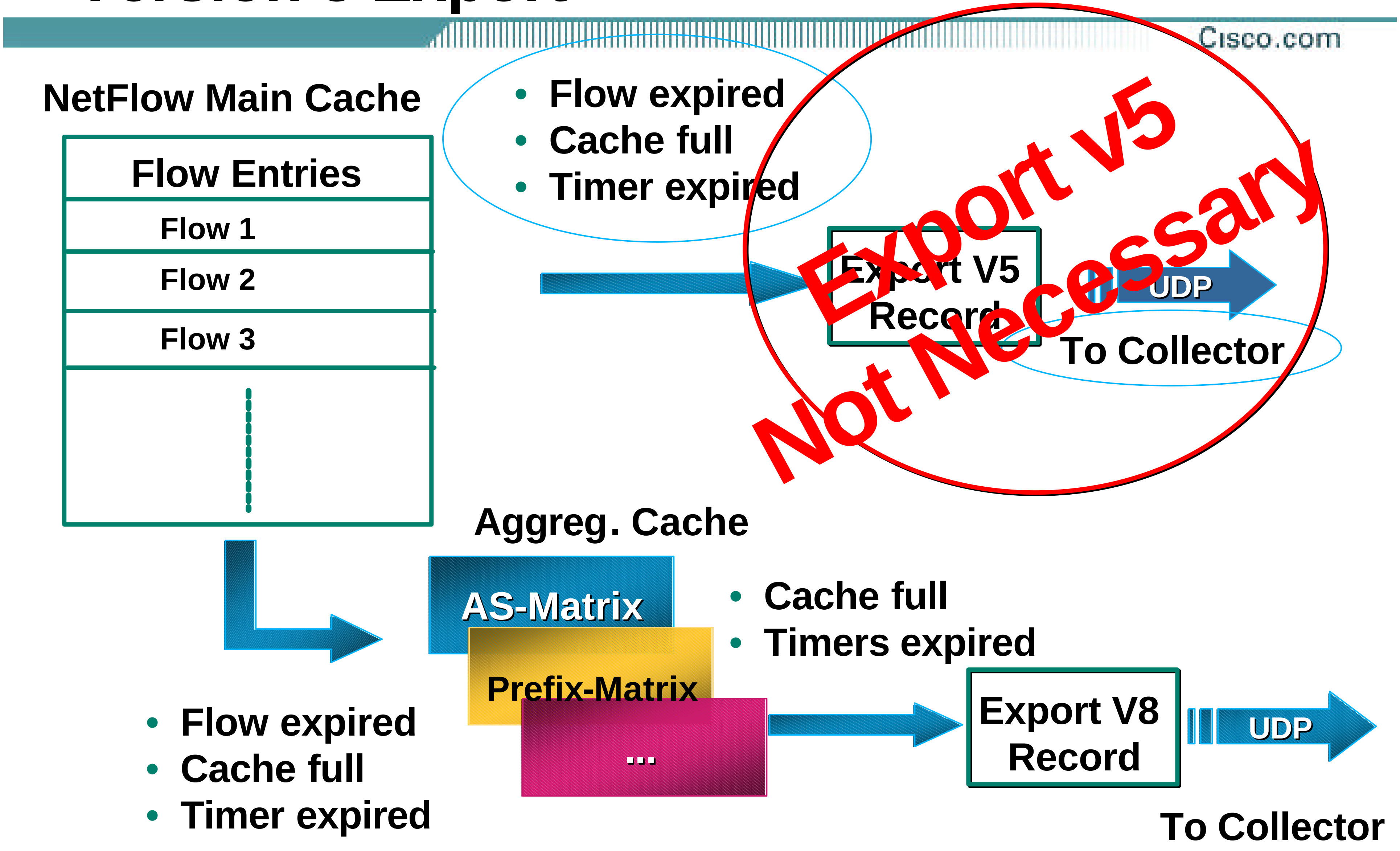
Version 8 - Flow Format

	AS	Protocol-Port	Source-Prefix	Destination-Prefix	Prefix
Source Prefix			·		·
Source Prefix Mask			·		·
Destination Prefix				·	·
Destination Prefix Mask				·	·
Source App Port		·			
Destination App Port		·			
Input Interface	·		·		·
Output Interface	·			·	·
IP Protocol		·			
Source AS	·		·		·
Destination AS	·			·	·
First Timestamp	·	·	·	·	·
Last Timestamp	·	·	·	·	·
# of Flows	·	·	·	·	·
# of Packets	·	·	·	·	·
# of Bytes	·	·	·	·	·

Version 8 - Flow Format

	AS-TOS	Protocol-Port-TOS	Source-Prefix-TOS	Destination-Prefix-TOS	Prefix-TOS	Prefix-Port
Source Prefix			·		·	·
Source Prefix Mask			·		·	·
Destination Prefix				·	·	·
Destination Prefix Mask				·	·	·
Source App Port		·				·
Destination App Port		·				·
Input Interface	·	·	·		·	·
Output Interface	·	·		·	·	·
IP Protocol		·				·
Source AS	·		·		·	
Destination AS	·			·	·	
TOS	·	·	·	·	·	·
First Timestamp	·	·	·	·	·	·
Last Timestamp	·	·	·	·	·	·
# of Flows	·	·	·	·	·	·
# of Packets	·	·	·	·	·	·
# of Bytes	·	·	·	·	·	·

Version 8 Export



NetFlow Main Cache

Flow Entries

Flow 1

Flow 2

Flow 3

- Flow expired
- Cache full
- Timer expired

Export V5 Record

UDP

To Collector

Export v5 Not Necessary

Aggreg. Cache

AS-Matrix

Prefix-Matrix

...

- Cache full
- Timers expired

- Flow expired
- Cache full
- Timer expired

Export V8 Record

UDP

To Collector

Version 8 Configuration

```
router (config)# ip flow-aggregation cache as

router (config-flow-cache)# export destination
  172.17.246.225 9996

router (config-flow-cache)# enabled




router (config)# ip flow-aggregation cache protocol-port

router (config-flow-cache)# export destination
  172.17.246.240 9996

router (config-flow-cache)# cache entries 8192

router (config-flow-cache)# enabled
```

Note the 2 different export ip addresses/ports

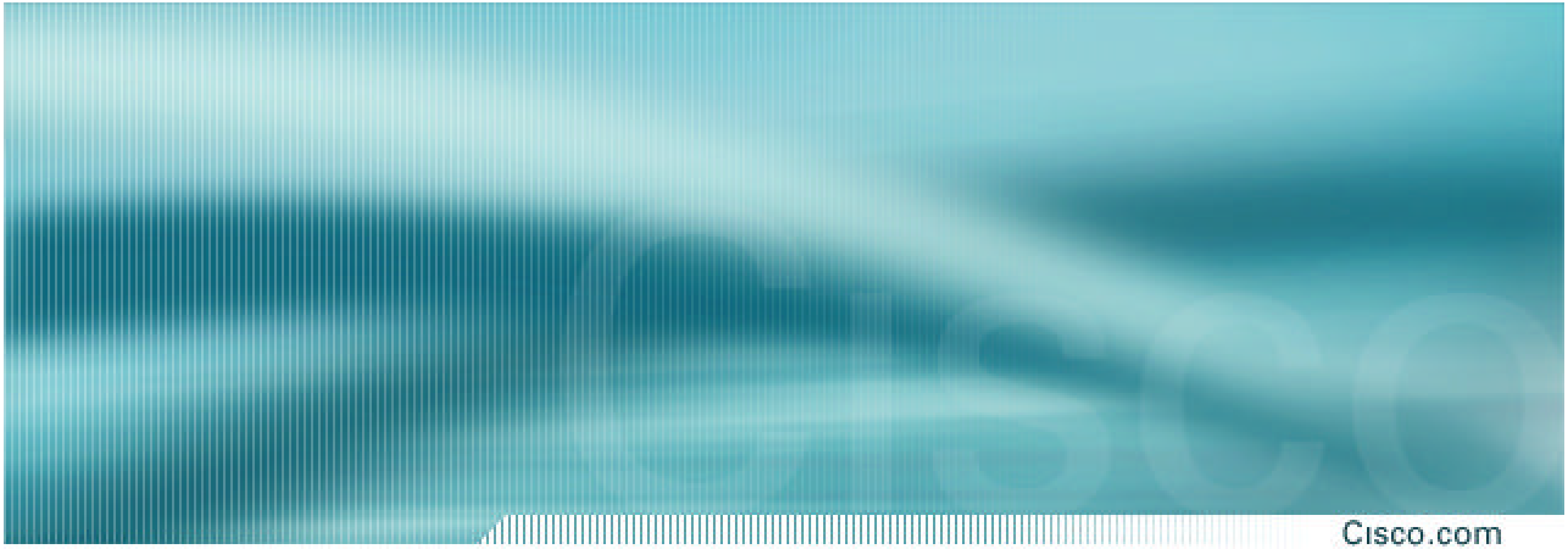
Version 8 Show Command

```
router#sh ip cache flow aggregation as

IP Flow Switching Cache, 278528 bytes  2 active, 4094
inactive, 13 added  216 ager polls, 0 flow alloc
failures

SrcIf    SrcAS  DstIf    DstAS Flows Pkts  B/Pk Active

Se0/0     0     Se0/2.1  0        1     1     104    0.0

Se0/0     0     Null     0        1     1     59     0.0
```

Note: you must choose peer-as or origin-as

```
router (config)# ip flow-export version 5 <peer-as
origin-as>
```

So that the main cache populates the BGP AS
So that the aggregation cache will contain the populated BGP AS



NetFlow on the 12000 Router
Sampled NetFlow

12000 NetFlow Sampling

- Collects and exports NetFlow data for a sample of the traffic passing through the router, instead of the entire traffic
- Only for the 12000 router (GSR) so far
- Sampled NetFlow exports the same information as full NetFlow
- The sampling interval is fixed and not an average
- Sampling advantages: CPU reduced and possible reduced exported Data
- Sampling disadvantage: no billing possible?

12000 NetFlow Sampling

```
Router(config)#ip flow-sampling-mode packet-interval
<10-16382>
Router(config-if)#ip route-cache flow sampled
```

Show Command

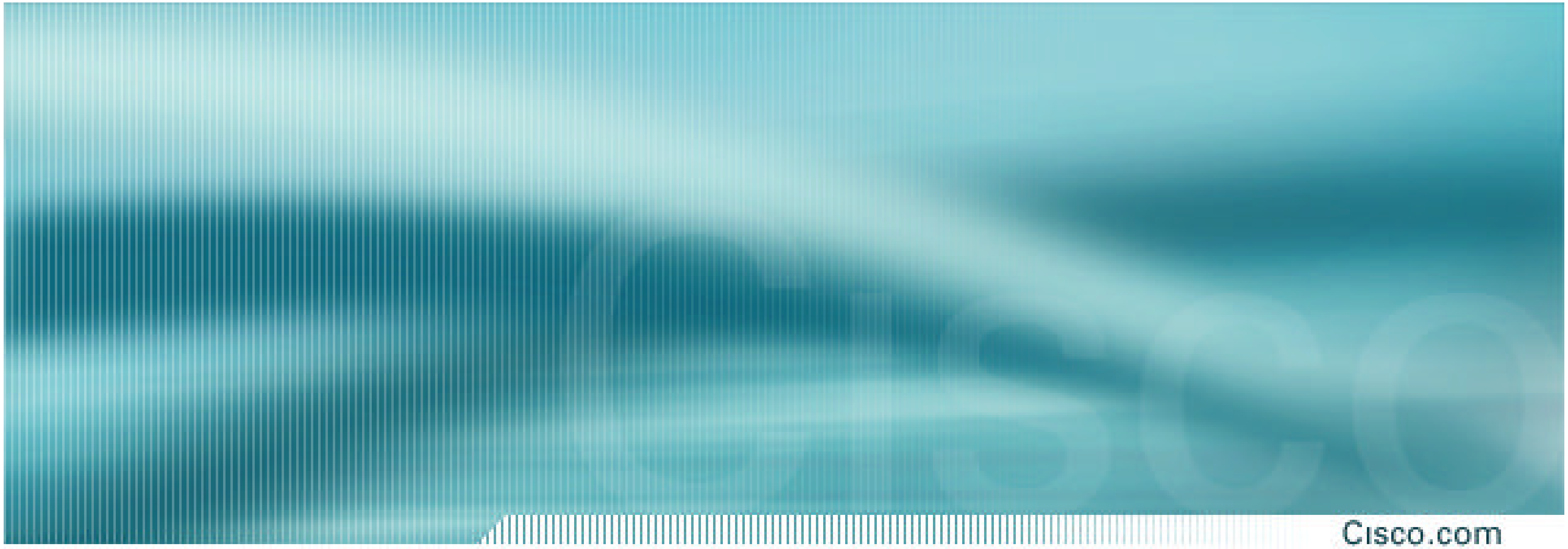
```
Router#show ip flow sampling
Flow sampling is enabled
'Packet Interval' sampling mode is configured.
1 out of every 100 packets is being sampled.
```

Status of NetFlow on the 12000 Series

		NetFlow		Sampled NetFlow	
		v5	v8	v5	v8
Engine 0		12.0(14)S	12.0(6)S	12.0(14)S	12.0(11)S
Engine 2	PoS	N/A	N/A	12.0(14)S	12.0(14)S
	3xGE	N/A	N/A	12.0(16)S	12.0(16)S
Engine 3		N/A	12.0(21)S	12.0(21)S	12.0(21)S
Engine 4		N/A	N/A	N/A	N/A
Engine 4+	PoS	N/A	N/A	12.0(21)S	12.0(21)S

Full NetFlow version 8
Engine 3 Line Cards

- No concept of main cache for full NetFlow version 8, the flows are directly created into the aggregation cache(s)
- Full NetFlow version 8 could be the solution versus Sampled NetFlow:
  - No main cache (the flow maintenance is the bottleneck)
  - Less flow in the aggregations cache
  - Export less flow
- Same behavior for the future engine 5 Line Cards



Advanced Concepts

Cache size

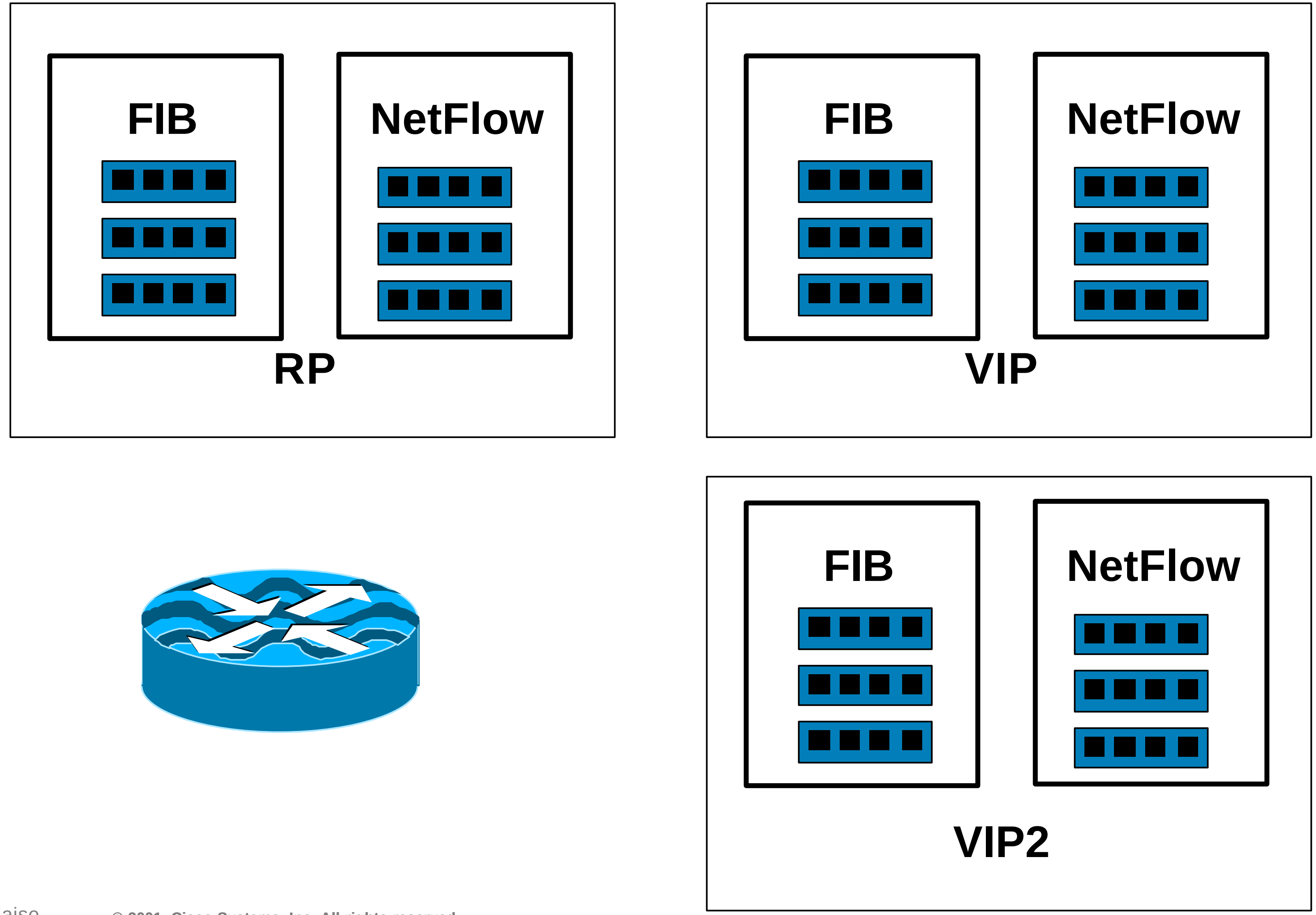
Platform	Default Netflow Cache Size (entries)	Approximate amount of contiguous DRAM used by Netflow cache
7x00, uBR7246, RSP7000	64K	4MB
AS5800, 4x00, 3600, 2600, 2500, 1600, 1400	4K	256KB
VIP with 128MB DRAM	128K	8MB
VIP with 64MB DRAM	64K	4MB
VIP with 32MB DRAM	32K	2MB
VIP with 16MB DRAM	2K	128K

Note that the latest IOS images don’t require contiguous DRAM anymore

12000 Line Card Cache size

Platform	Default Netflow Cache Size (entries)	Approximate amount of contiguous DRAM used by Netflow cache
LC with 1024MB DRAM	1M	64MB
LC with 512MB DRAM	512K	32MB
LC with 256MB DRAM	256K	16MB
LC with 128MB DRAM	128K	8MB
LC with 64MB DRAM	64K	4MB
LC with 32MB DRAM	32K	2MB
LC with 16MB DRAM	8K	512kB

Version 5 VIP/LC caches



FIB

NetFlow

RP

FIB

NetFlow

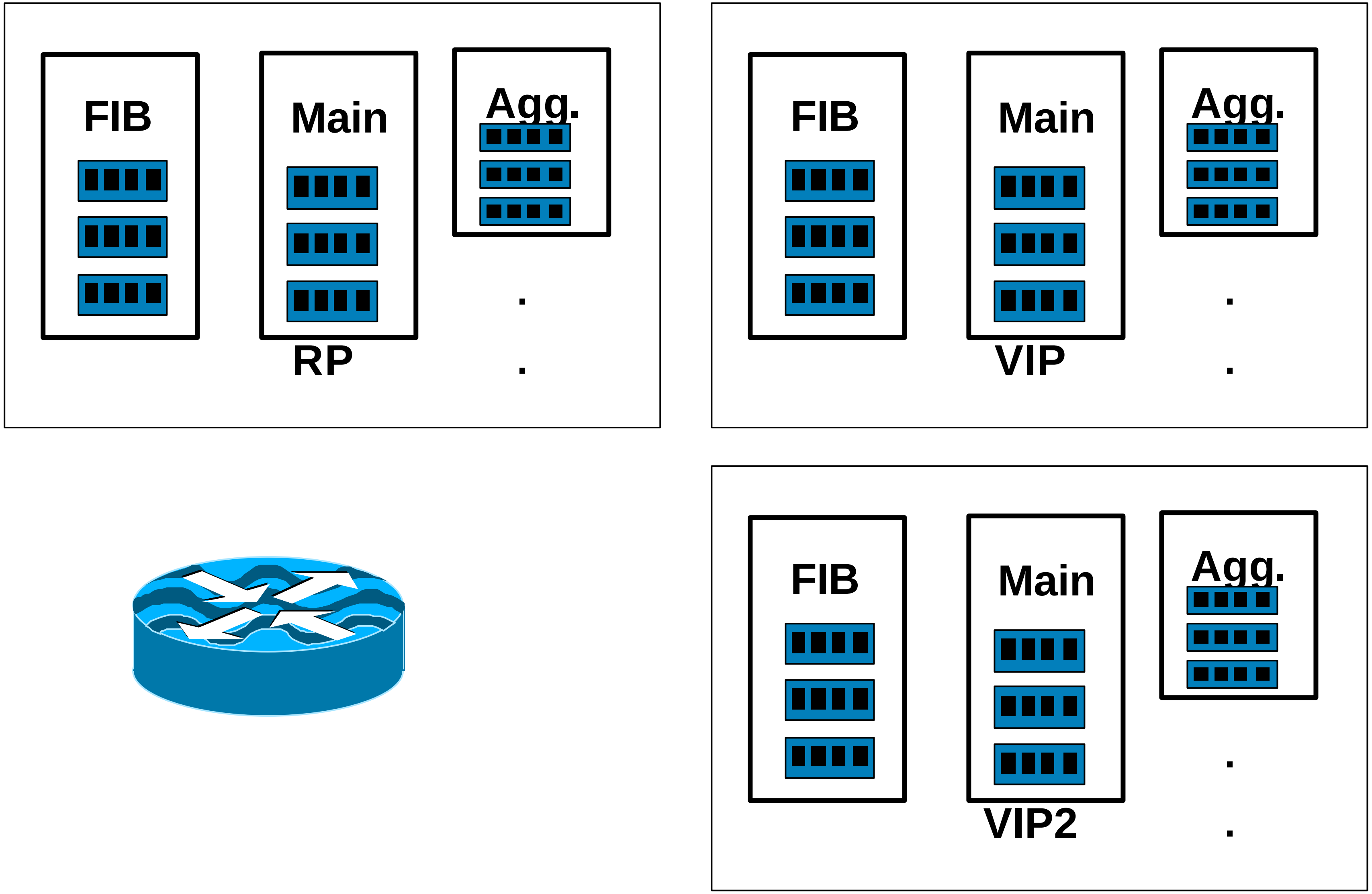
VIP

FIB

NetFlow

VIP2

Version 8 VIP/LC Caches



VIP/LC Caches

- Nothing to configure on the VIP/LC (use DCEF)
- VIP:

```
if-con <slot-number>
        sh ip cache flow
```

- LC:

```
attach <slot-number>
        sh ip cache flow
Execute-on <slot-number> show …
```

- Own independent sequence numbering per VIP/LC
- Note: Don’t export on the engine management ethernet port on the 12000, even though it’s a possible configuration

Flow Ageing

- **When is a flow expired?**

  Transport is completed (TCP FIN or RST)

  After 15 sec of traffic inactivity (the only way for UDP). The inactive timer

  After 30 min of traffic activity. The active timer.

  The cache is becoming full

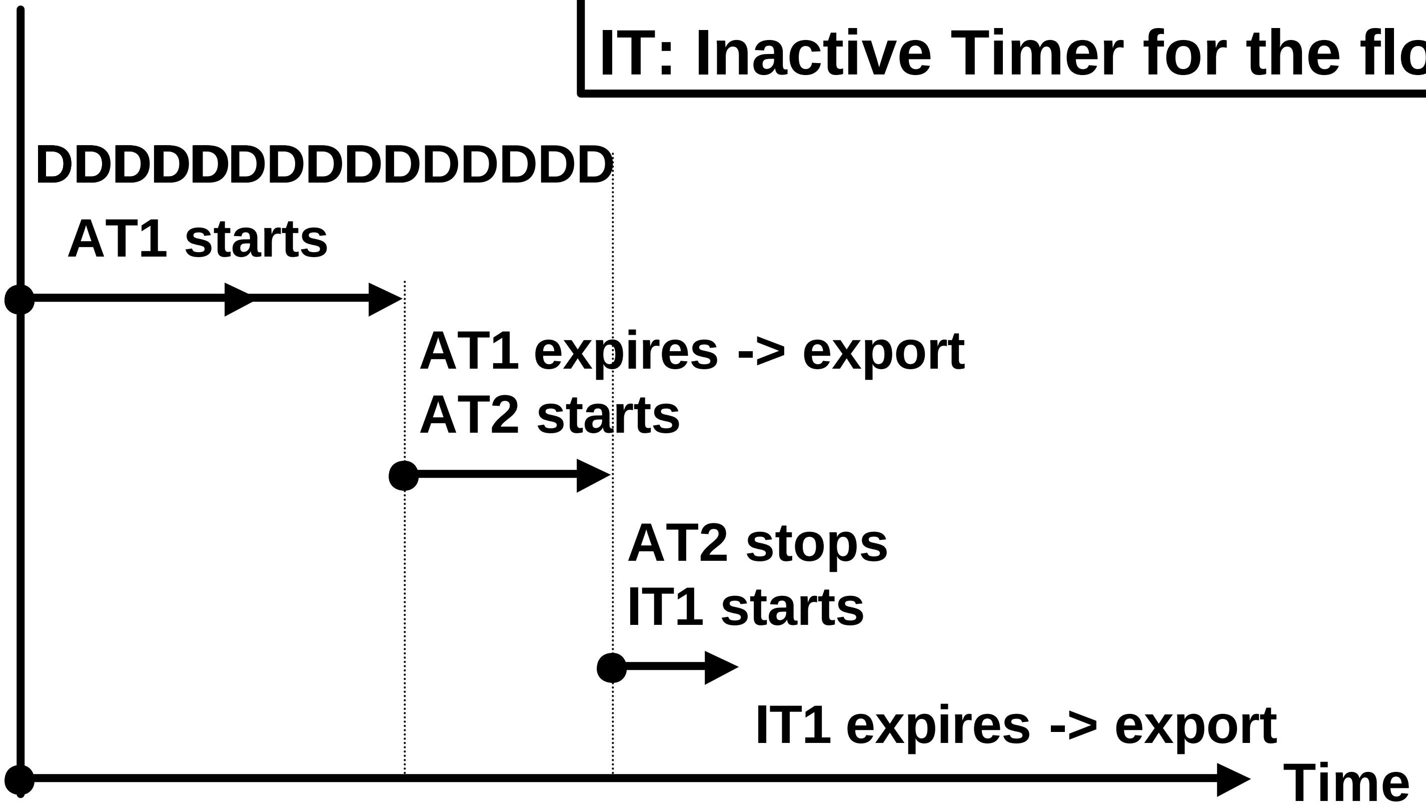
  Note that 15sec/30min are the router default timers

Active/Inactive Timers

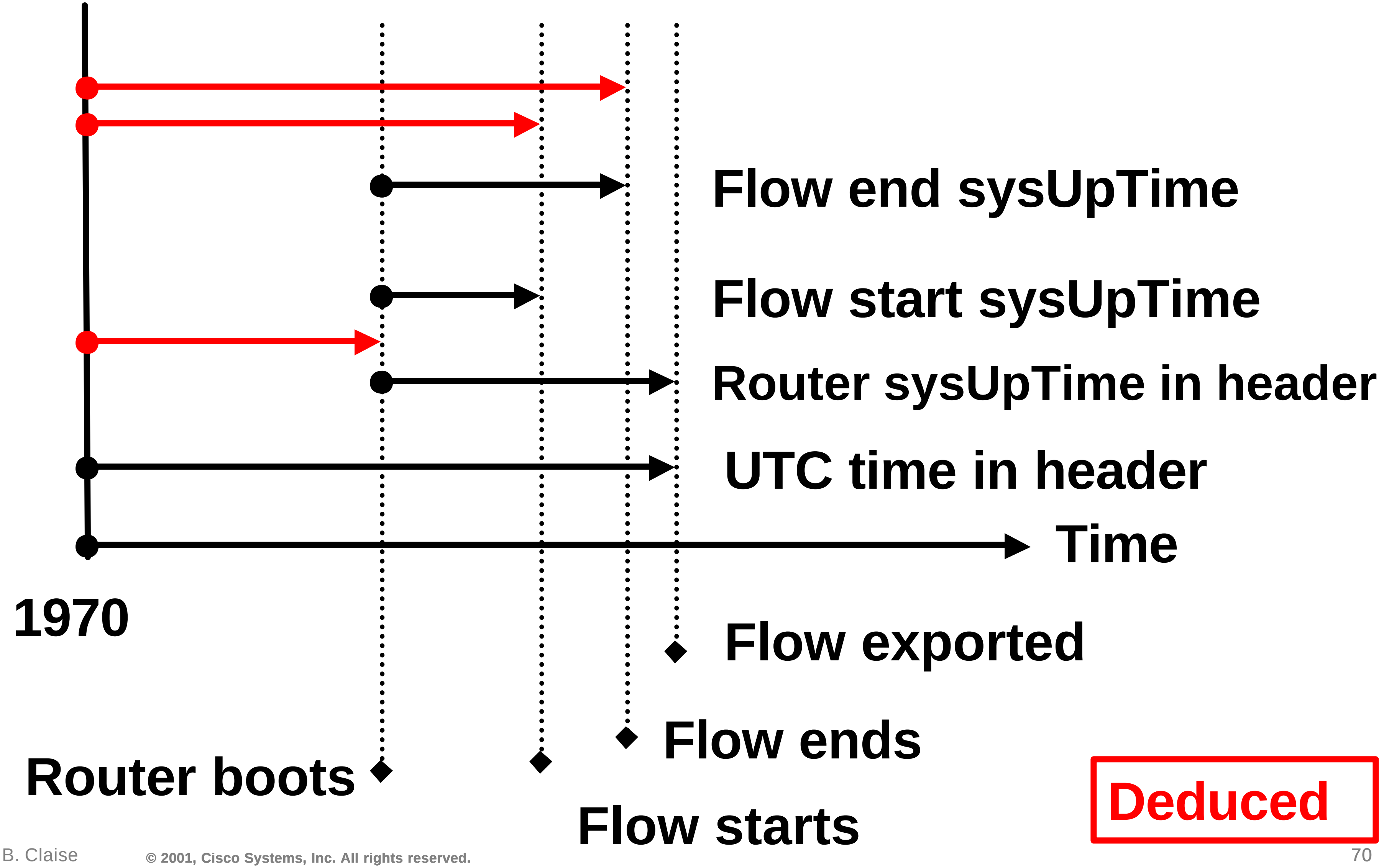
D: Data (UDP)

AT: Active Timer for the flow

IT: Inactive Timer for the flow



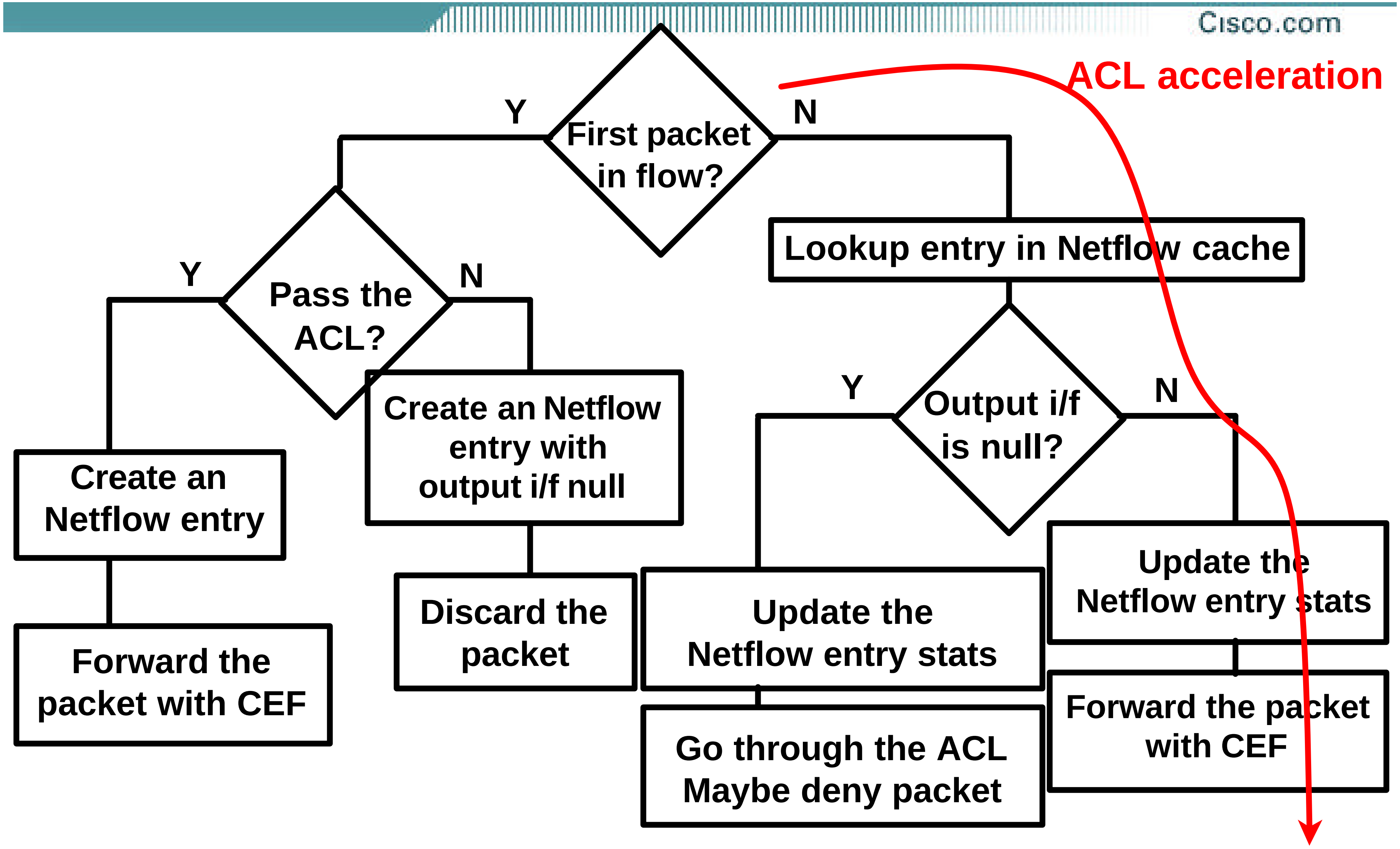
Various Time in NetFlow



Various Time in NetFlow

- The UTC depends on the clock
- Synchronization of the VIP clock, the line card clock (in sync. since 12.0) and the RSM/MSFC clock
- Attention to the timezone on the collector
- Conclusion: the device clocks must be synchronized
- NTP is a solution, NTP MIB in 12.1(4)

NetFlow Bypasses the Access-list

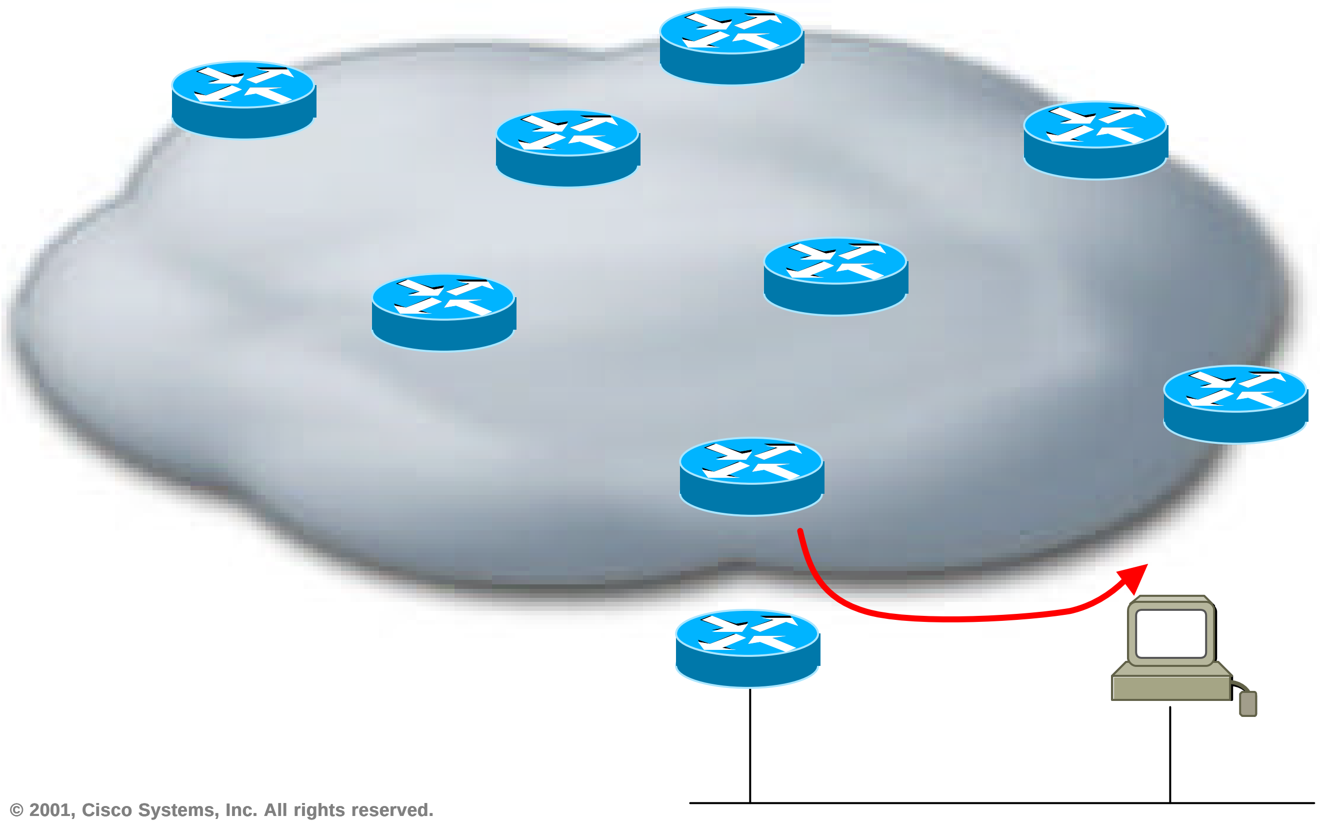


ACL acceleration

First packet in flow?
Y
N
Pass the ACL?
Y
N
Create an Netflow entry
Forward the packet with CEF
Create an Netflow entry with output i/f null
Discard the packet
Lookup entry in Netflow cache
Output i/f is null?
Y
N
Update the Netflow entry stats
Go through the ACL Maybe deny packet
Update the Netflow entry stats
Forward the packet with CEF

NetFlow and DOS attack

Sh ip cache verbose <server ip address> flow



Performance (Approximate Number)

- Enabling NetFlow version 5 AND exporting increases the cpu utilization by around 15 % (with a max of 20 % depending on the platform)
- Enabling Neflow version 8 increases the cpu utilization by 2 to 5%, depending on the number of aggregations enabled
  With a multiple of 6% for multiple aggregations
- NetFlow is done in hardware on the cat6000 supervisor and the 12000 Engine 3 Line Cards

NetFlow Performance testing: Results at a Glance

CPU impact:

10,000 active flows: < 4% of additional CPU utilization

45,000 active flows: <12% of additional CPU utilization

65,000 active flows: <16% of additional CPU utilization

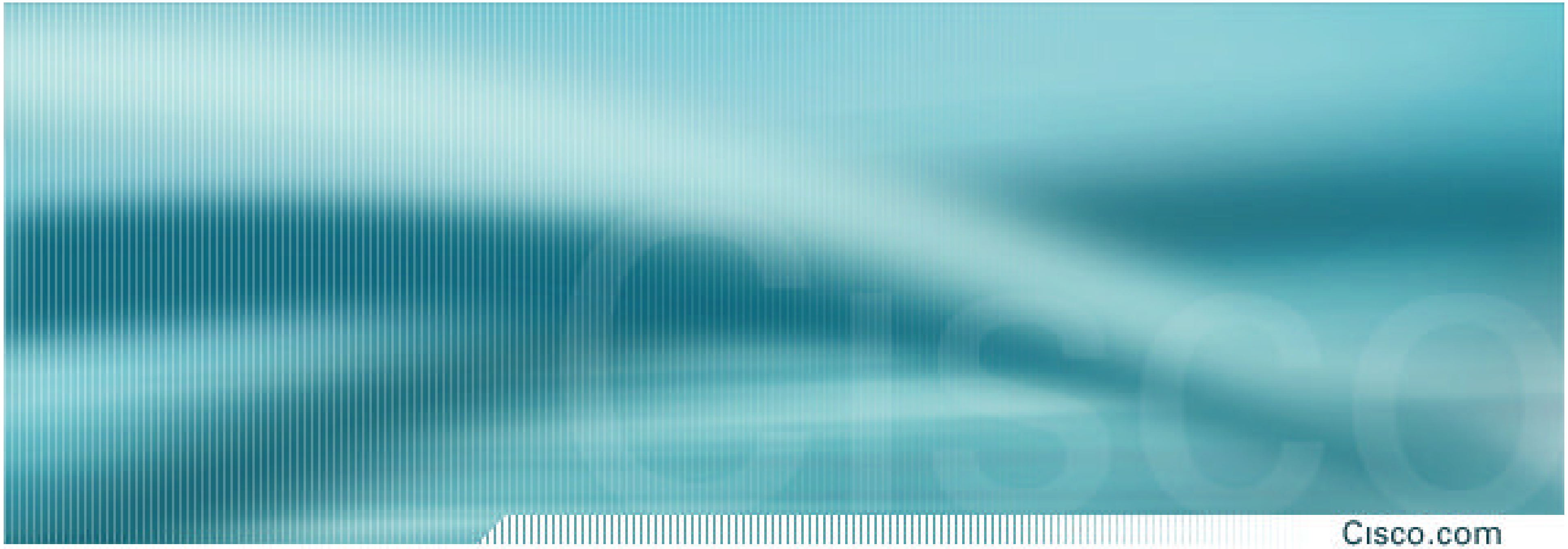
NetFlow Data Export (single/dual): no real impact

NetFlow v5 vs. v8: minimal to no impact at all

NetFlow Feature Acceleration: >200 lines of ACLs

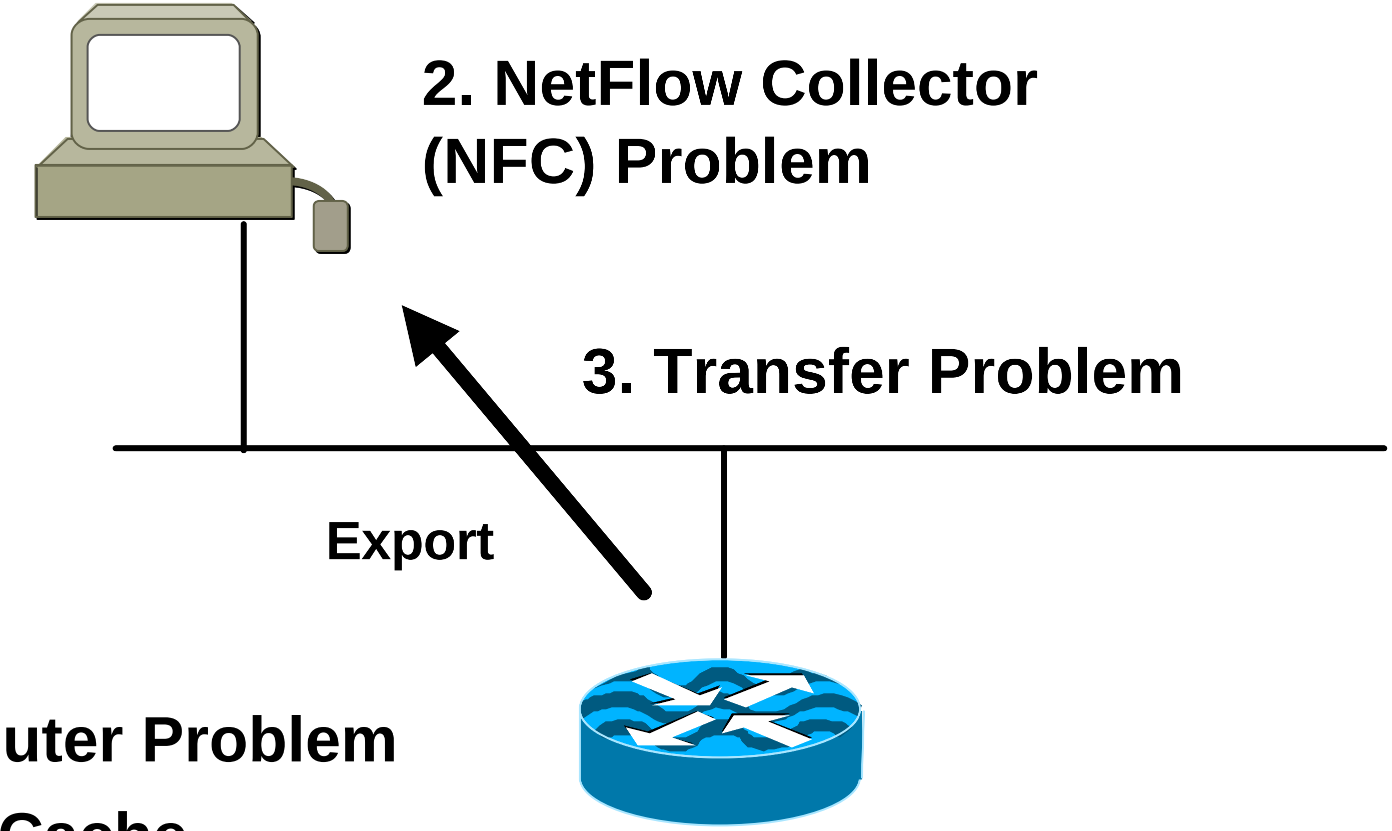
Sampled NetFlow on the Cisco 12000:

23 % vs 3 % (65,000 flows, 1:100)



Troubleshooting

Missing Flows?



2. NetFlow Collector (NFC) Problem

3. Transfer Problem

Export

1. Router Problem

- Cache
- Export

Missing Flows?
- 1. Router Problem

Cisco.com

```
Router#sh ip cache flow (excerpt)
     IP Flow Switching Cache, 4456704 bytes
      2 active, 65534 inactive, 226352 added
      3792086 ager polls, 0 flow alloc failures
      Active flows timeout in 40 minutes
      Inactive flows timeout in 20 seconds
      82038 flows exported in 34439 udp datagrams, 0 failed
      last clearing of statistics 00:14:23
```

Alloc failures: Number of times the NetFlow code tried to allocate a flow but could not

Failed: Number of flows that could not be exported by the router because of output interface limitations

Missing Flows?
- 1. Router Problem

```
Router#sh ip flow export
  Flow export is enabled
  Exporting flows to 151.99.57.3 (9996)
  Exporting using source interface Loopback0
  Version 5 flow records, origin-as
  2304658131 flows exported in 219987515 udp datagrams
  0 flows failed due to lack of export packet
  167 export packets were sent up to process level
  0 export packets were punted to the RP
  3490 export packets were dropped due to no fib
  7012 export packets were dropped due to adjacency issues
  0 export packets were dropped enqueuing for the RP
  0 export packets were dropped due to IPC rate limiting
  0 export packets were dropped due to output drops
```

Missing Flows?
- 2. NFC Problem

- The Netflow Collector “show tech-support”

```
udpPort: 9996, receivedFlows: 80277(0),
receivedFlowrecords: 1771469(0)

discardedFlows: 0, missedFlowrecords:
1115(0),  socNum: 13, rcvQSize: 26000
```

Missing Flows?
- 2. NFC Problem

- **Netstat -s**

```
udpInDatagrams = 14034  udpInErrors =  0
udpInCksumErrs = 0 udpInOverflows =3218
```

- **In Netflow Collector, the number of missed records is directly proportional to the number of rules and the order of rules.**

```
Filter deny-traffic-x
  Deny    Srcaddr      24.192.1.19     0.0.0.0
```

Missing Flows?
- 3. Transfer Problem

- The only remaining explanation
- Don’t forget that the NetFlow exported data are transported over UDP
- Evaluate the exported traffic

Exported Traffic Estimation

- Rule of thumb:

  Export 1 % to 1.5% of the total box throughput

- To be more accurate, you need:

  packet/sec of throughput (router figures, sh int switching)

  Ex: 150kpps average throughput on a 7500

  average number of packets per flow
  (sh ip cache flow)

  Ex: 20 (a number recently quoted for Internet backbone traffic)

Exported Traffic Estimation

- **Example for a 7500:**

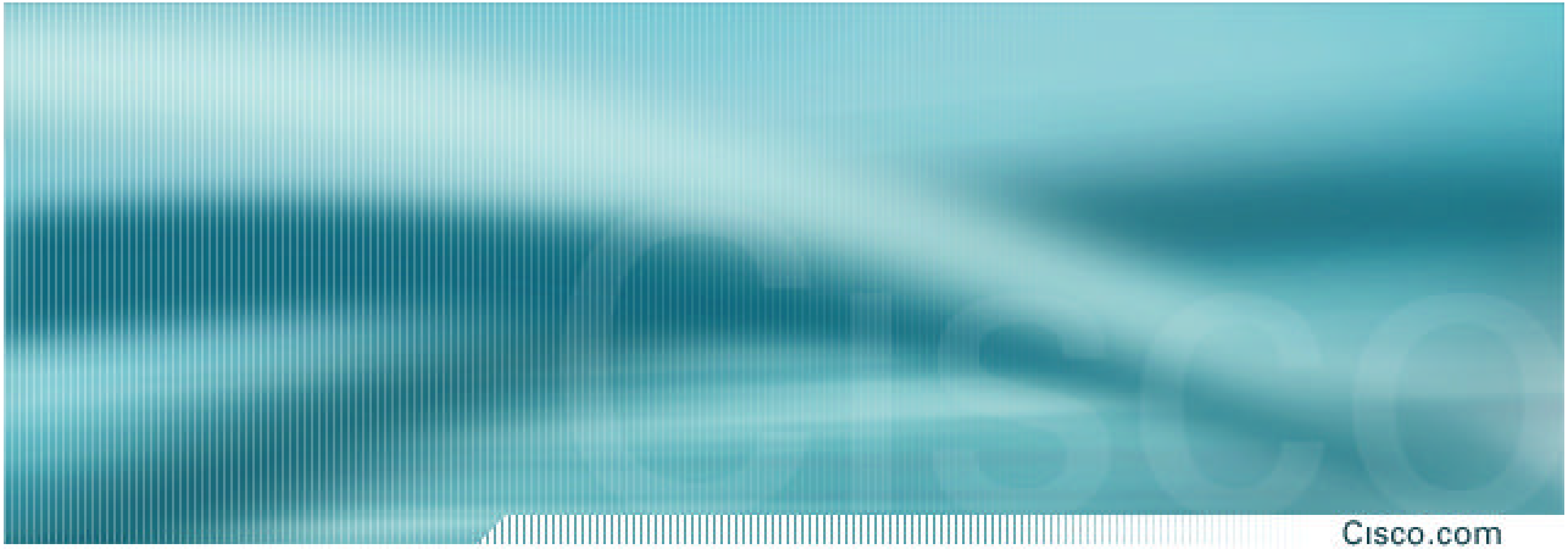
150kpps / 20 ppflow = 7500 flow / sec

Considering 30 flows per exported packet and a length of 1500 bytes

7500 /30 *1500 = 375 Kbytes/sec of flow export traffic from one router

Flows/Packet

	Number of flow in a packet	Packet length (bytes)
V1	24	Approx. 1200
V5	30	Approx. 1500
V7	28	Approx. 1500
V8 AsMatrix	51	1456
V8 ProtocolPortMatrix	51	1456
V8 SourcePrefixMatrix	44	1436
V8 DestinationPrefixMatrix	44	1436
V8 PrefixMatrix	35	1428



New Features

ifIndex Persistence

- No guarantee that the ifIndex values for any “interface” will remain the same after a reboot.
- The NetFlow exports contain the input/output interfaces ifIndex
- Introduced in 12.0(11)S, 12.0(11)SC and 12.1(5)T

```
router(conf) snmp-server ifindex persist
router(conf-if) snmp-server ifindex persist
```

NetFlow on Egress for MPLS Traffic

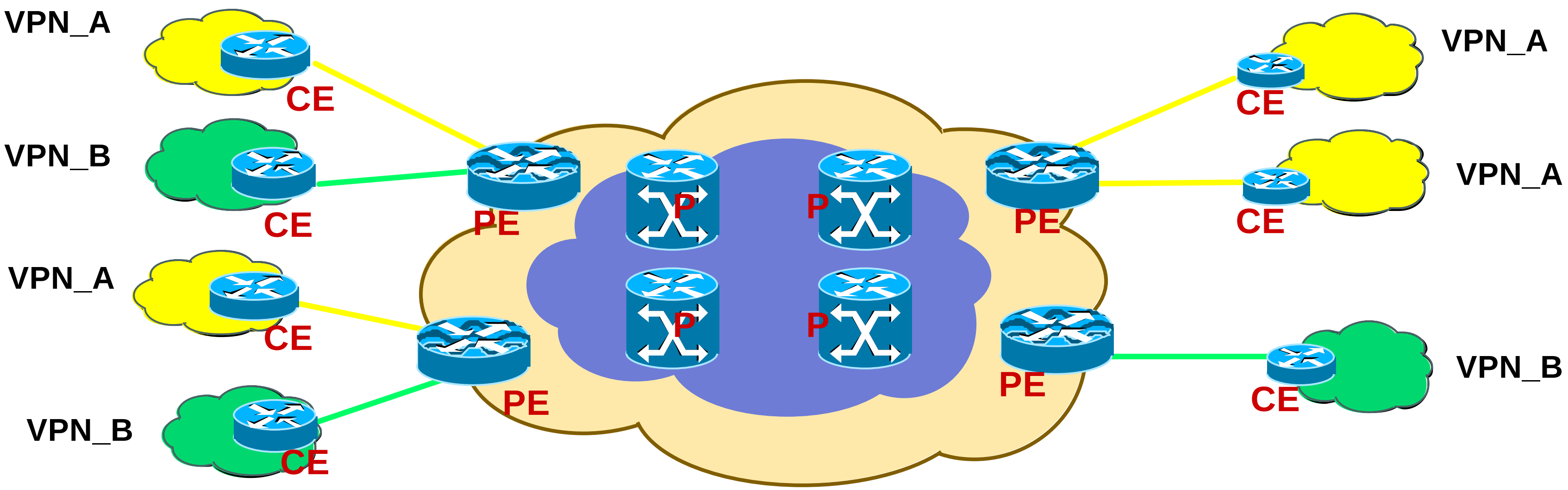
New

- Introduced in 12.0(10)ST, 12.1(5)T, 12.0(22)S
- For MPLS/VPN traffic only, i.e. the traffic coming from the core
- Caches traffic on the egress interface, not the ingress interface.
- Valid for version 5 and version 8

```
router(config-if)#tag-switching ip flow egress
```

- Can be enabled on subinterface
- All other NetFlow commands still apply

NetFlow on Egress for MPLS Traffic



- Now: enable egress/ingress on one PE
- Can deduce the packets lost in the core
- No accounting if both src and dst VPNs are part of the same PE

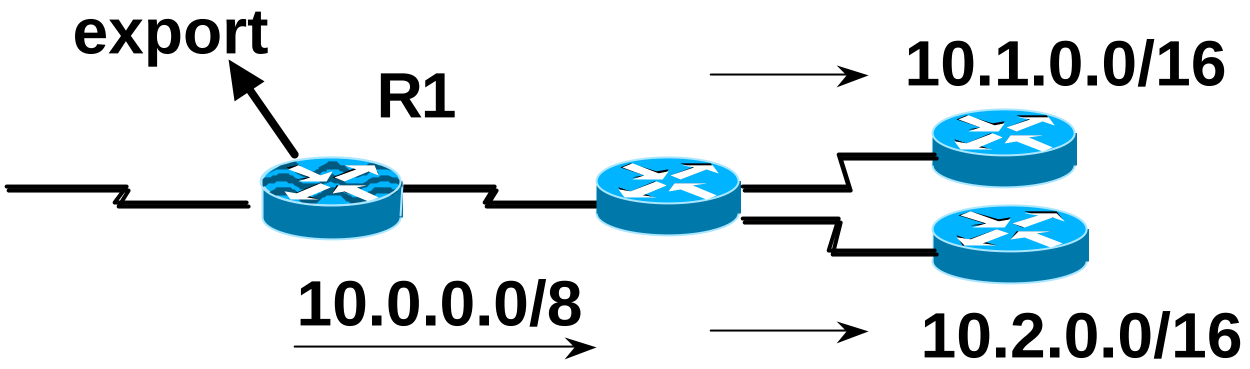
Minimum Prefix Mask for Router-Based Aggregation

	AS	Protocol-Port	Source-Prefix	Destination-Prefix	Prefix
Source Prefix			·		·
Source Prefix Mask			·		·
Destination Prefix				·	·
Destination Prefix Mask				·	·

- Prefixes come from the routing table
- Introduced in 12.0(11)S, 12.1(2)T
- Only for the Aggregations:

  SourcePrefix, DestinationPrefix and Prefix

Minimum Prefix Mask for Router-Based Aggregation



- **Summarization on the router R1**
- **Lose the granularity unless we specify the minimum mask of 16**

Minimum Prefix Mask for Router-Based Aggregation

- Configuration:

```
router (config)# ip flow-aggregation cache prefix
router (config-flow-cache)# mask source minimum 24
router (config-flow-cache)# mask destination minimum 16
```

- SourcePrefix: only source
- DestinationPrefix: only destination

Dual Flow Export

New

- Inserted into 12.2(2)T, 12.0(19)S and 12.0(19)ST, 2 redundant export destinations are allowed for version 5

```
router(config)#ip flow-export destination 1.1.1.1 9996
router(config)#ip flow-export destination 2.2.2.2 9997
```

If try to configure more, you will get:

“Exceeded maximum export destinations”

- Only for the routers, not the catalysts for now

Cat6000 Aggregations – Version 8

New

- Add 3 new aggregation schemes: RouterDestOnly, RouterSrcDst, RouterFullFlow
- Hybrid version since CatOS version 5.5(2)
  Not on Native version yet
- Must select the nde version 8 instead of 7
- Require the NetFlow Collector 3.6 or above
- No real aggregations (like version 8 on routers)
  - Because still IP addresses and no networks
  - The aggregation is defined by the flow mask

Cat6000 Aggregations – Version 8

	RouterDstOnly	RouterSrcDst	RouterFullFlow
Source IP address		·	·
Destination IP address	·	·	·
Source App Port			·
Destination App Port			·
IP Protocol			·
First Timestamp	·	·	·
Last Timestamp	·	·	·
# of Flows	·	·	·
# of Packets	·	·	·
# of Bytes	·	·	·

No real aggregation like on a router, where we aggregate IP addresses in prefixes

Cat6x00 Switched Traffic

New

- The switched type traffic (intra vlan) is now accounted with NetFlow
- Since CatOS version 7.(2)
  Not on Native version yet

```
“set mls bridged-flow-statistics enable/disable
  <vlan>“
```

Cat6x00 New Fields Population

New

- **SUP2/PFC2 (EARL6) supports from 12.1(13)E:**

  Source and Destination BGP AS

  Input and Output ifIndexes

  Next Hop

- **Note: 12.1(13)E1 if any WAN cards**

Cat6x00 NetFlow Version 5 Support

New

- SUP2/PFC2 supports NetFlow version 5 from 12.1(13)E
- Some consistency…

NetFlow on Subinterface

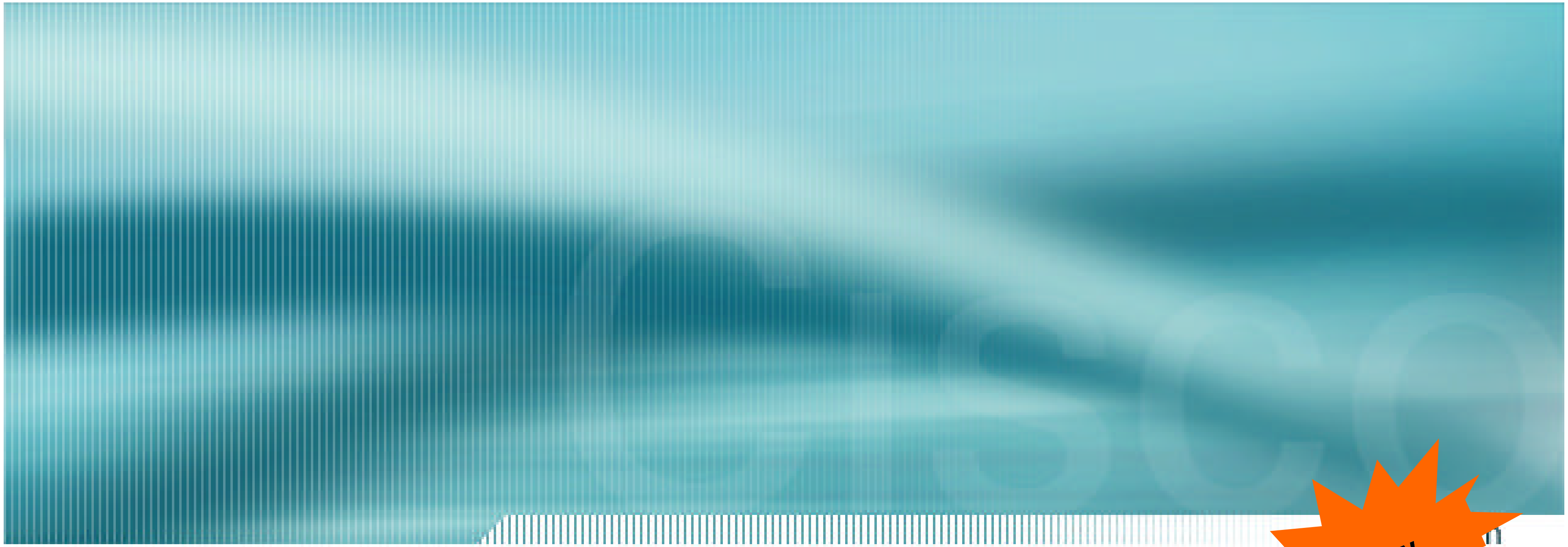
New

- Introduced in 12.0(21)S
- Under investigation for the 12000

Egress Sampled NetFlow

New

- Egress Sampled NetFlow on engine 3
- IP->IP and MPLS->IP cases
- Available 12.0(24)S, for the 12000



New

New Features
NetFlow Version 9 and IETF

NetFlow Version 9
Why do we need a New Version?

- Fixed formats for export
  - Easy to implement
  - Consume little bandwidth
  - Easy to decipher at the collector
- But
  - Not flexible and not extensible
- Consequence
  - Always new aggregations for new combinations of fields and for new technologies required
  - New collector versions required each time

Version 9 Approach

- Current NetFlow versions are not flexible and not extensible
- Version 9 based on template and separate flow record

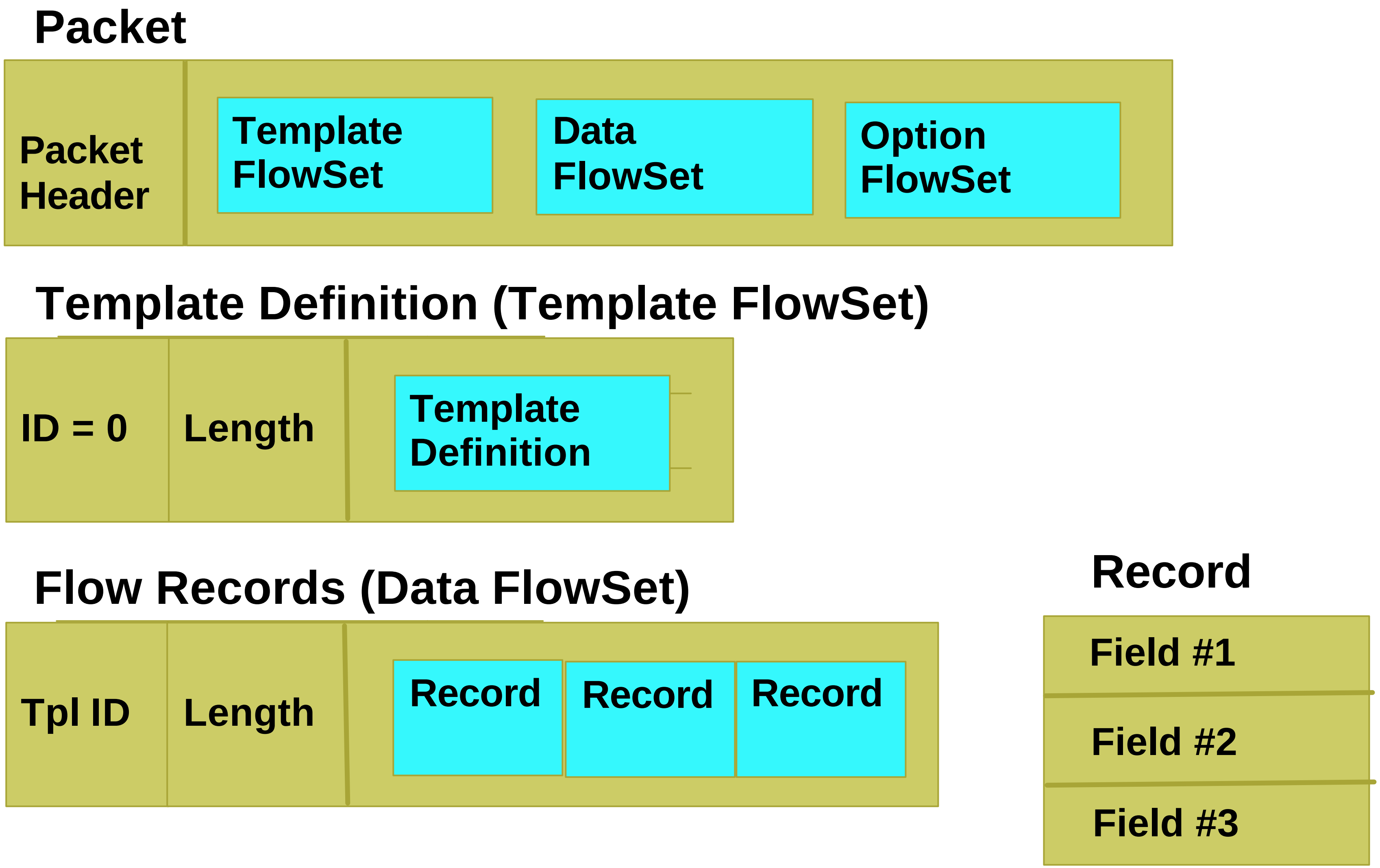
  Template composed of type and length

  Flow record composed of template ID and value

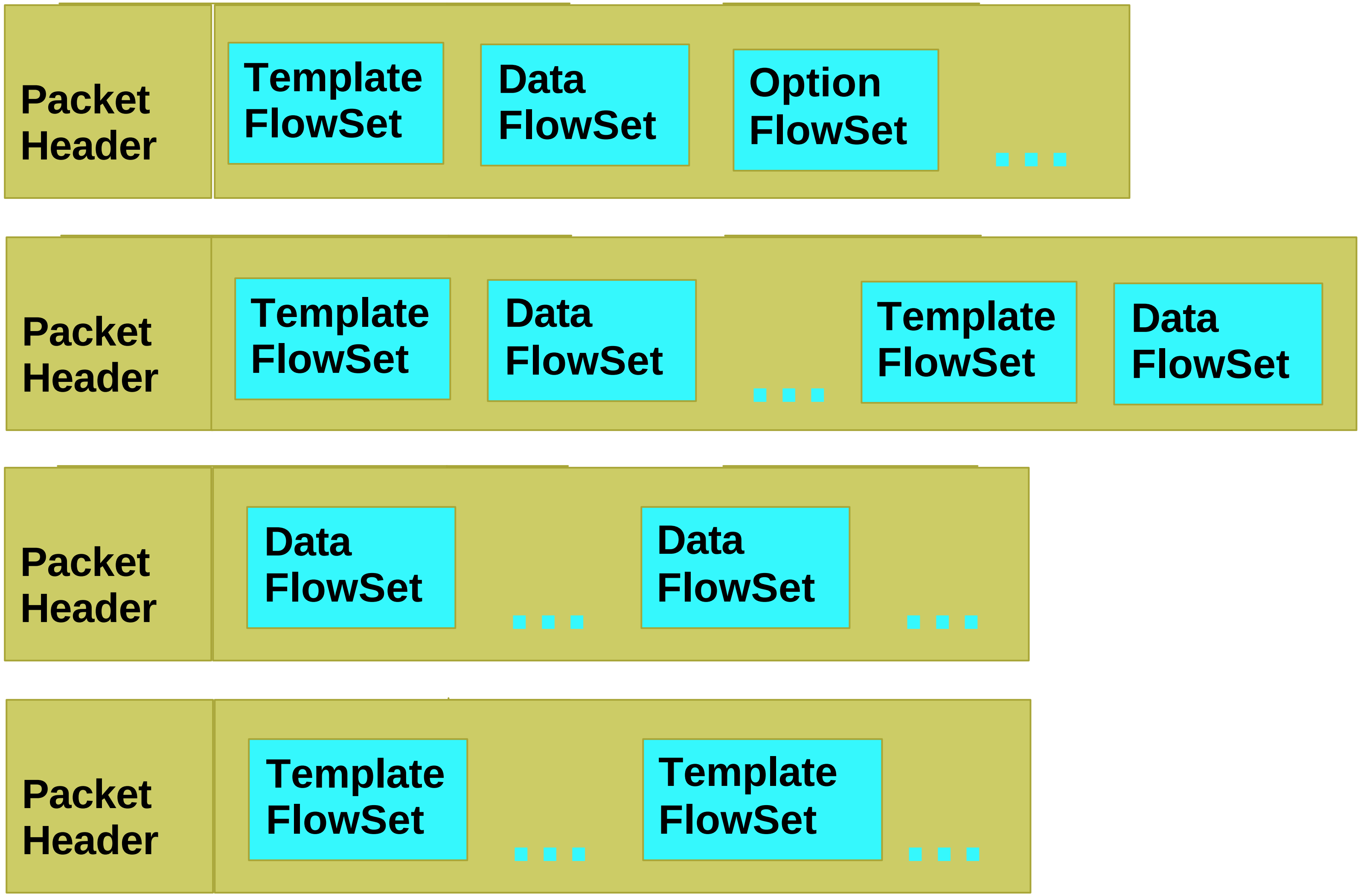
- Whitepaper

  http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/tflow_wp.htm

NetFlow Version 9



NetFlow Version 9
Various Type of Export Packets



Packet Header | Template FlowSet | Data FlowSet | Option FlowSet | . . .

Packet Header | Template FlowSet | Data FlowSet | . . . | Template FlowSet | Data FlowSet

Packet Header | Data FlowSet | . . . | Data FlowSet | . . .

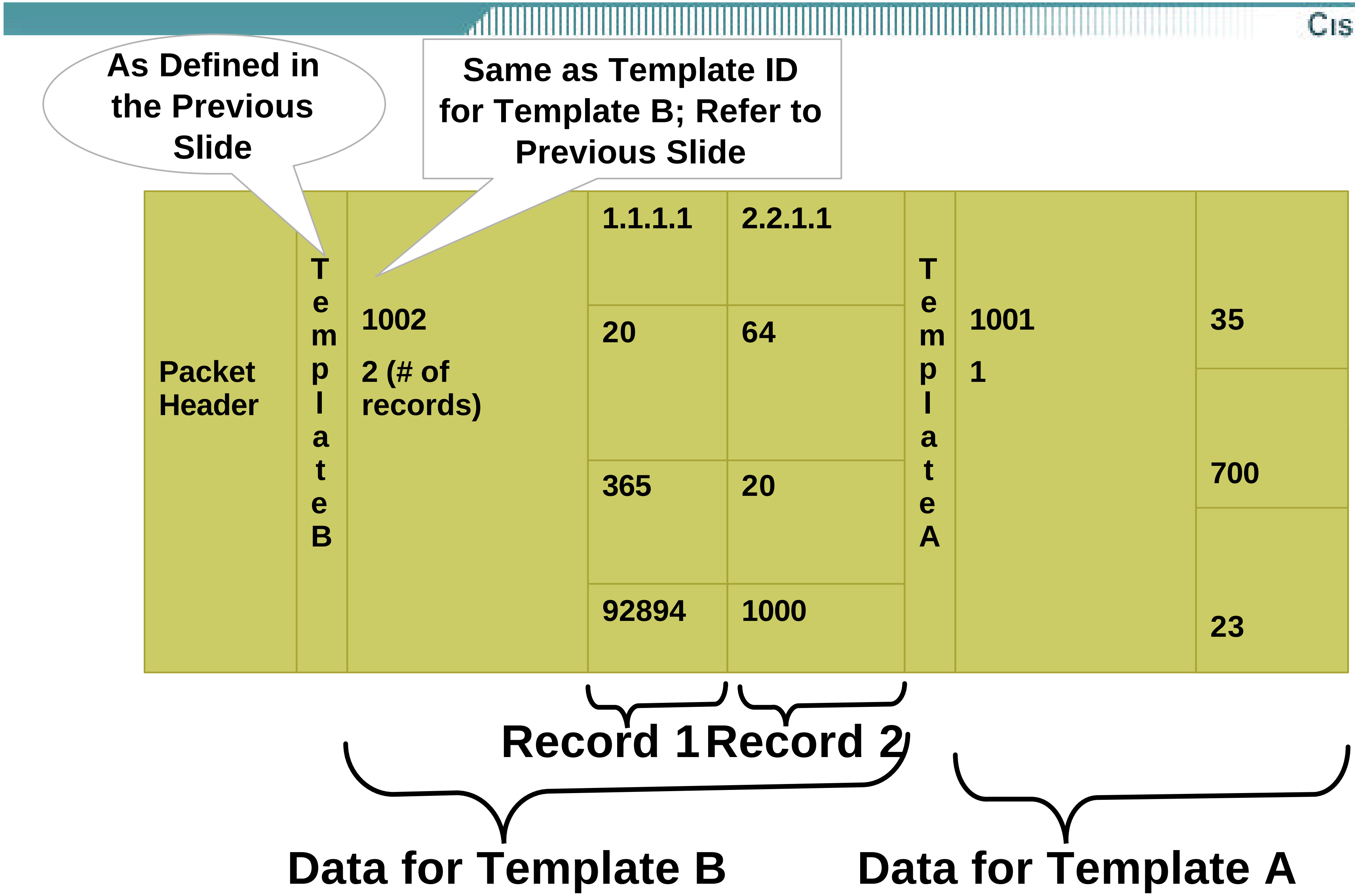
Packet Header | Template FlowSet | . . . | Template FlowSet | . . .

Version 9
Example for Template Definition

Template A
Flow Set ID (0 for Template)
Length of Template Structure
1001 (Template ID)
3 (# of Fields)
SRC_AS_NUMBER
2
DST_AS_NUMBER
2
L4_PROTOCOL
2

Template B
Flow Set ID (0 for Template)
Length of Template Structure
1002 (Template ID)
4 (# of Fields)
SRC_IP_PREFIX
4
SRC_AS_NUMBER
2
PACKET_COUNT
2
BYTE_COUNT
2

Example for Export Packet



As Defined in the Previous Slide

Same as Template ID for Template B; Refer to Previous Slide

Packet Header

Template B

1002

2 (# of records)

1.1.1.1

20

365

92894

2.2.1.1

64

20

1000

Template A

1001

1

35

700

23

Record 1

Record 2

Data for Template B

Data for Template A

NetFlow version 9 Principles

- Still a push model
- Sent the template regularly (configurable)
- Independent of the underlying protocol, ready for any reliable protocol (thinking of SCTP)
- FlowSet Flexibility in the export packet

NetFlow version 9 Support

- Out in 12.0(24)S
- Committed for 12.3T
- Cafeteria based aggregation on the router is not yet available

IETF: IP Flow Information Export WG (IPFIX)

- Internet Protocol Flow Information eXport (IPFIX) is an effort to standardize flow export
- IPFIX web site for the charter, email archive, drafts, etc. http://ipfix.doit.wisc.edu/
- Cisco's NetFlow version 9 has been presented a the first BOF
- Cisco actively participating, authors of the 3 current drafts

IPFIX Working Group at IETF

- **Requirements draft:** **http://www.ietf.org/internet-drafts/draft-ietf-ipfix-reqs-08.txt**
- **Architecture draft:** **http://www.ietf.org/internet-drafts/draft-ietf-ipfix-architecture-01.txt**
- **Data Model draft:** **http://www.ietf.org/internet-drafts/draft-ietf-ipfix-data-00.txt**

Version 9 and IPFIX

- Cisco NetFlow Version 9 draft: http://www.ietf.org/internet-drafts/draft-bclaise-netflow-9-00.txt

  Next version will become an I-RFC

- “Intellectual Property Rights” Notice on the IETF web site because there is a patent for NetFlow

IPFIX Next Steps

- The requirement draft will go “last call” pretty soon
- An evaluation team is created:
  - Evaluation existing protocols: NetFlow, CRANE, LFAP, Diameter, IPDR
  - Choose THE base protocol
  - Determine which improvements are needed for THE protocol compared to the requirements
- Hopefully, NetFlow will be chosen

NetFlow and the IPFIX Evaluation

- draft-claise-ipfix-eval-netflow-03.txt
- NetFlow compliant to most of the points
- Biggest exceptions:

  MUST run on the top of a congestion aware export protocol

  MUST have authenticity, integrity, SHOULD have confidentiality

New

New Features
MPLS aware NetFlow Solution

MPLS aware NetFlow Description

- Provides flow statistics per MPLS and IP packets
  - MPLS packets:
    - Labels information
    - And the V5 fields of the underlying IP packet
  - IP packets:
    - Regular IP NetFlow records
- Based on the NetFlow version 9 export
- Configure on ingress interface
- Supported on sampled/non sampled NetFlow

NetFlow MPLS Aware Support

- Supported in 12.0(24)S, then 12.2S and maybe 12.2T

  Support on the 12000: Engine 0, 1, 2, 3 and 4+

- Will be supported on 12.0(26)S on the 7200/7500
- The catalyst 6000 will only support the export of the top label, due to hardware limitations

NetFlow MPLS Aware Flow Keys

- Key Fields (Uniquely Identifies the flow)
  - Source IP address
  - Destination IP address
  - IP Protocol
  - Input ifIndex
  - Source Application Port
  - Destination Application Port
  - DSCP
  - Up to 3 incoming MPLS labels of interest with experimental bits and end-of-stack bit
  - Positions of the above labels in the packet label stack

- Additional Export Fields
  - Flows
  - Packets
  - Bytes
  - First SysUptime
  - Last SysUptime
  - Output interface
  - NetFlow version 5 fields of the underlying IP packet
  - Type of the top label: LDP, BGP, VPN, ATOM, TE Tunnel MID-PT, unknow
  - The Forwarding Equivalent Class mapping to the top label

NetFlow MPLS Aware What is exported?

- Export up to 3 incoming MPLS labels
- Experimental bits and end-of-stack bit
- Positions of the above labels in the label stack
- Type of the top label:
  LDP, BGP, VPN, ATOM, TE Tunnel MID-PT, unknown
- The Forwarding Equivalent Class mapping to the top label, i. e. the IP address of the IBGP peer in a MPLS (VPN) environment

NetFlow MPLS Aware
What is exported?

- **Underlying IP packet: will export the NetFlow V5 fields of the underlying IP packet, when available:**

  Src and Dst AS, subnet masks and IGP next hop are not available! Null will be exported

- **Underlying non-IP packet: will export the NetFlow V5 fields:**

  Src and Dst IP addresses, protocol, TOS, application ports and TCP flags will be set to Null!

NetFlow MPLS Aware Configuration

Cisco.com

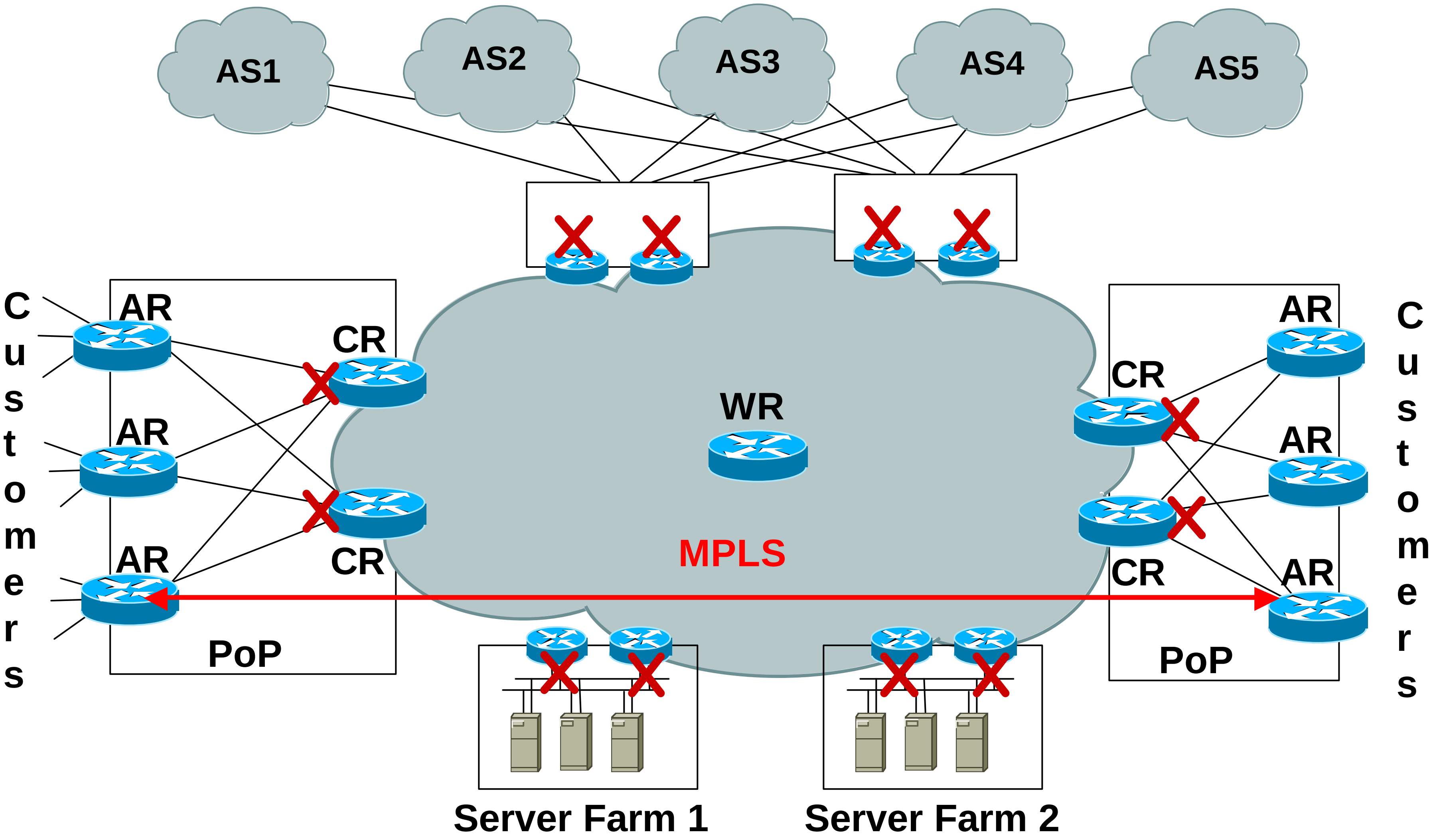
```
router (config)# ip flow export version 9
router (config)# ip flow-export template options sampling
router (config)# ip flow-export template options export_stats
router (config)# ip flow-export template options timeout 5
router (config)# ip flow-export template refresh-rate 10
router (config)# ip flow-sampling-mode packet-interval 101

router (config)# ip flow-cache mpls label-positions [1] [2] [3]
router (config-if)# ip route-cache flow sampled
```

Label position is starting from the top label, 1 corresponds to the top of the stack

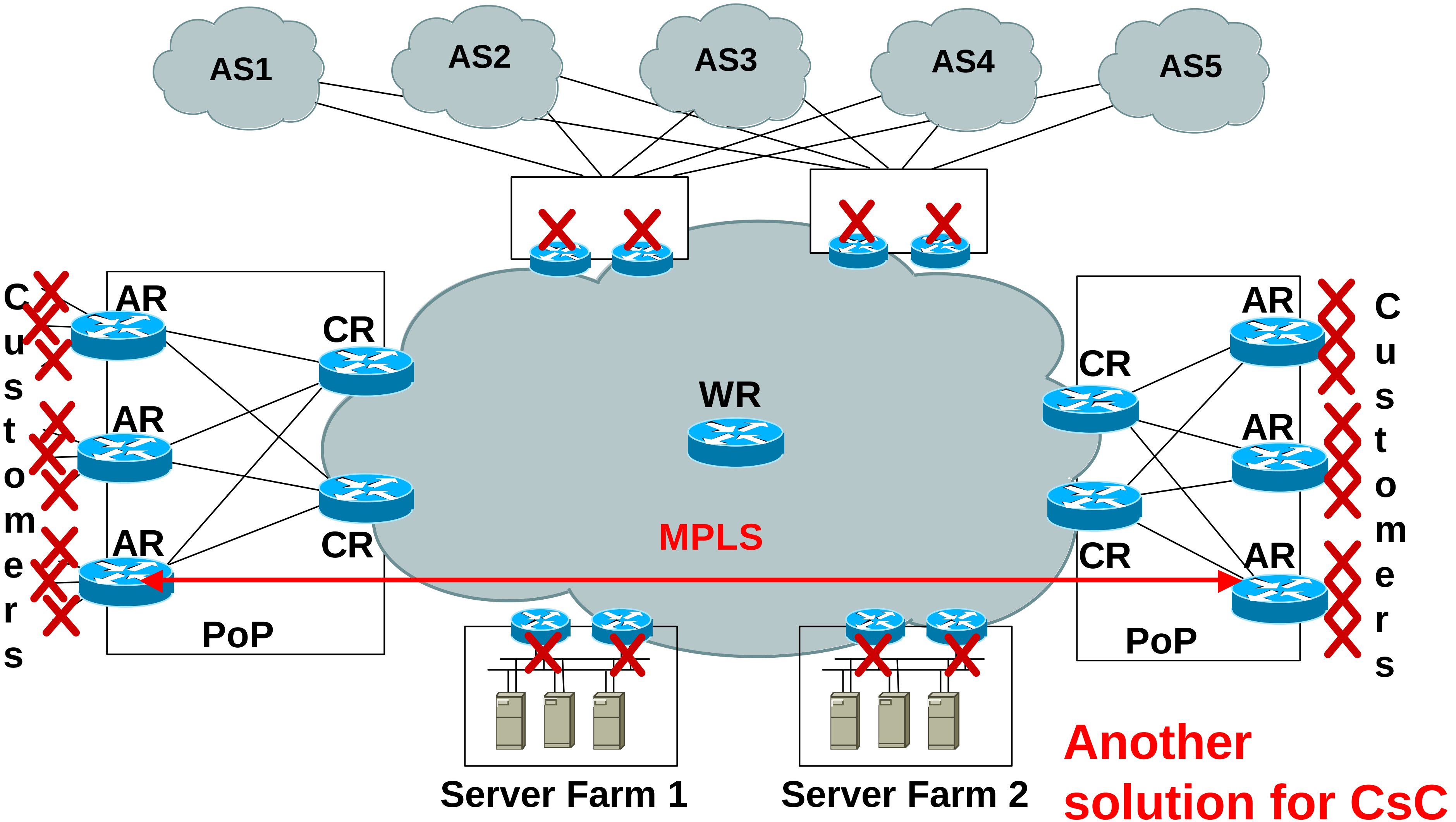
NetFlow MPLS Aware
Show commands

```
LC-Slot# show ip cache verbose flow
...
SrcIf        SrcIPaddress  DstIf        DstIPaddress  Pr TOS Flgs  Pkts
Port Msk AS                 Port Msk AS  NextHop       B/Pk   Active

PO1/0        8.1.1.1       PO4/0:1      80.0.0.1      06 00  00    24K
0100 /0  0                 0200 /0  0   0.0.0.0       256    34.6
Pos:Lbl-Exp-S 1:12305-6-0 (LDP/20.20.20.20) 2:12312-6-
```

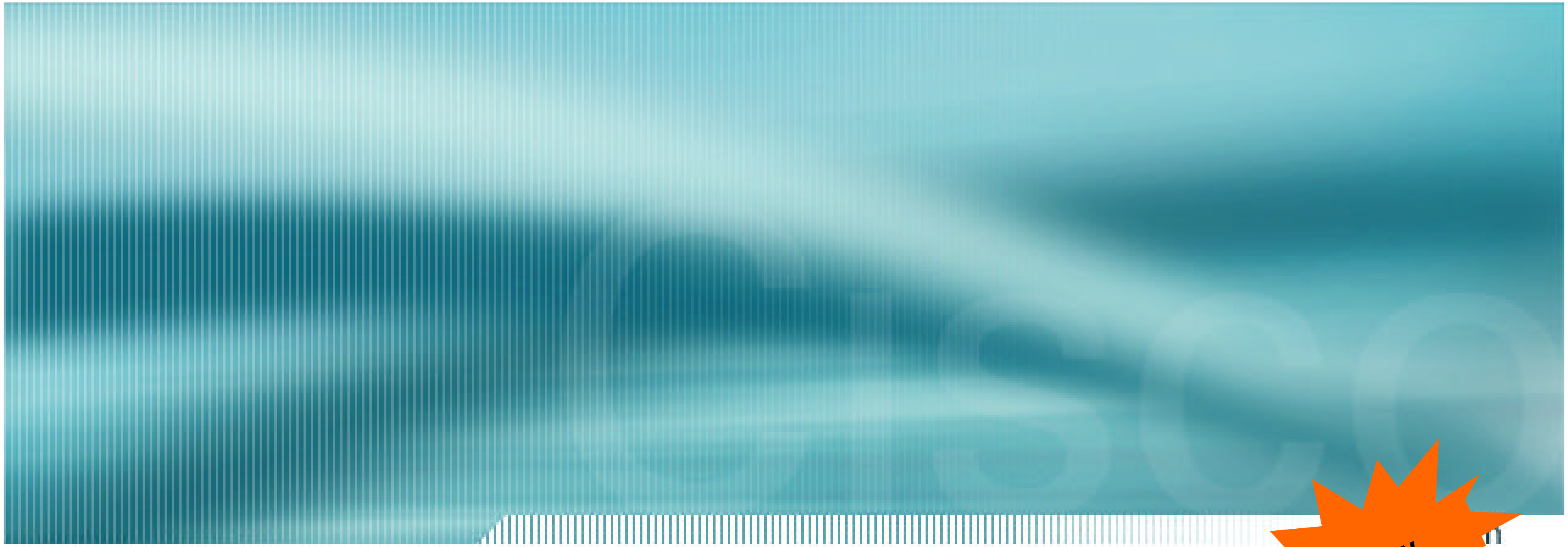
NetFlow MPLS Aware
Typical Example



NetFlow MPLS Aware
Typical Example



Another solution for CsC



New

New Features
BGP Next Hop TOS aggregation

NetFlow BGP Next Hop TOS Aggregation

- New NetFlow aggregation on the Router
- Only for the BGP routes
- For IP packets (not MPLS)
- Also available under the VRF interface
- Configure on ingress interface
- Take the BGP Next Hop from the “via” fields in “sh ip cef <destination_IP_address>”

NetFlow BGP Next Hop TOS Aggregation Support

- Currently on EFT

  Currently EFT, since September

- GSR will follow later:

  BGP next hop in 12.0(26)S

- Available on a wide range of platforms

  Initially 7200 & 7500 then 1720, 2600, 3600, 4500, 4700, 5800, RSP 7000, RSM (Cat5000), 7200, 7500, MGX Router Processor Module (RPM), 8800, GSR

NetFlow BGP Next Hop TOS Aggregation Flow Keys

- Key Fields (Uniquely Identifies the flow)
  - Origin AS
  - Destination AS
  - Inbound Interface
  - DSCP (*)
  - Next BGP Hop
  - Output Interface

- Additional Export Fields
  - Flows
  - Packets
  - Bytes
  - First SysUptime
  - Last SysUptime

(*) before any recoloring

Core Capacity Planning

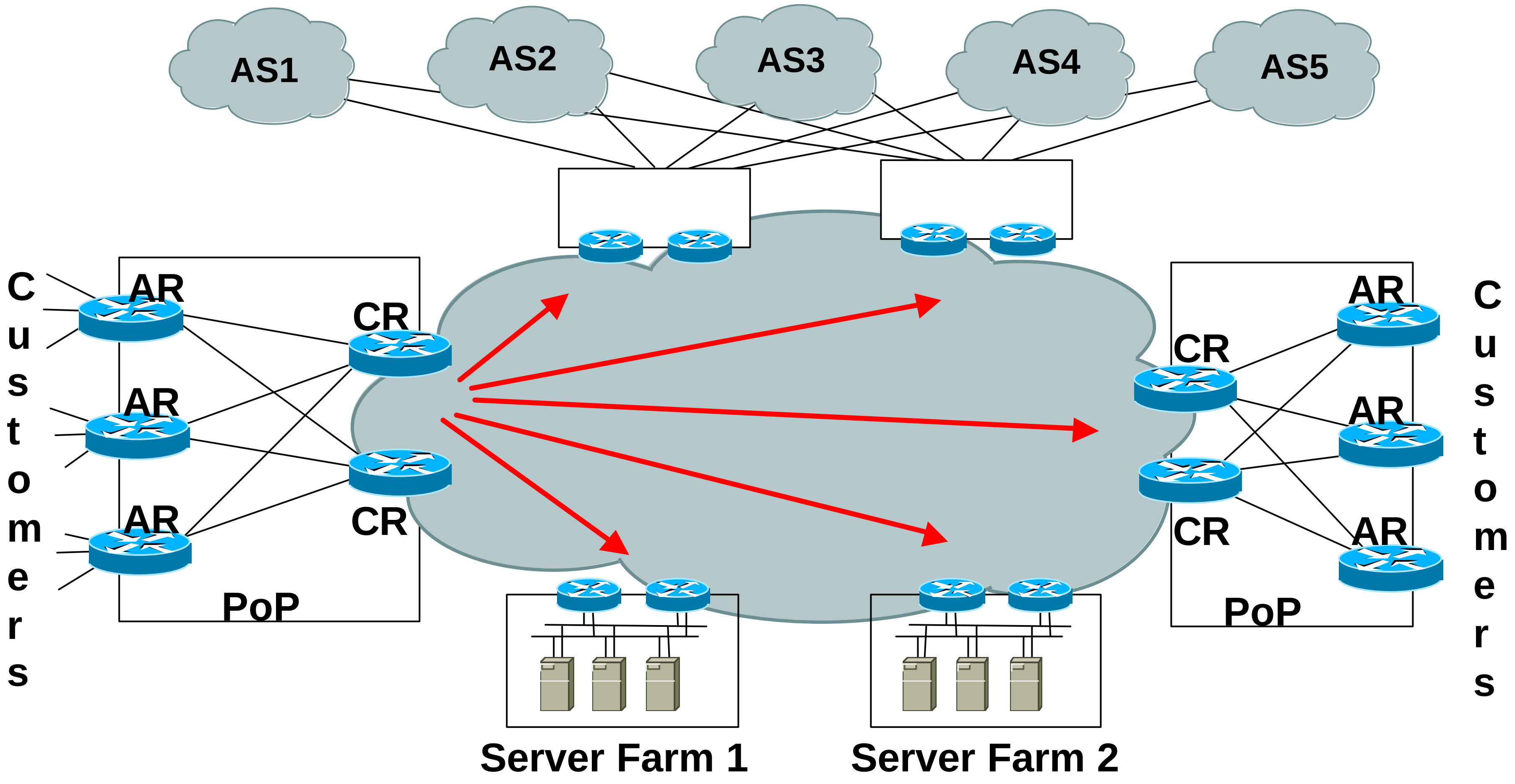
- The ability to offer SLAs is dependent upon ensuring that core network bandwidth is adequately provisioned
- Adequate provisioning (without gross over provisioning) is dependent upon accurate core capacity planning

Core Capacity Planning
What input?

- **Accurate core capacity planning is dependent upon understanding the core traffic matrix and flows and mapping these to the underlying topology**

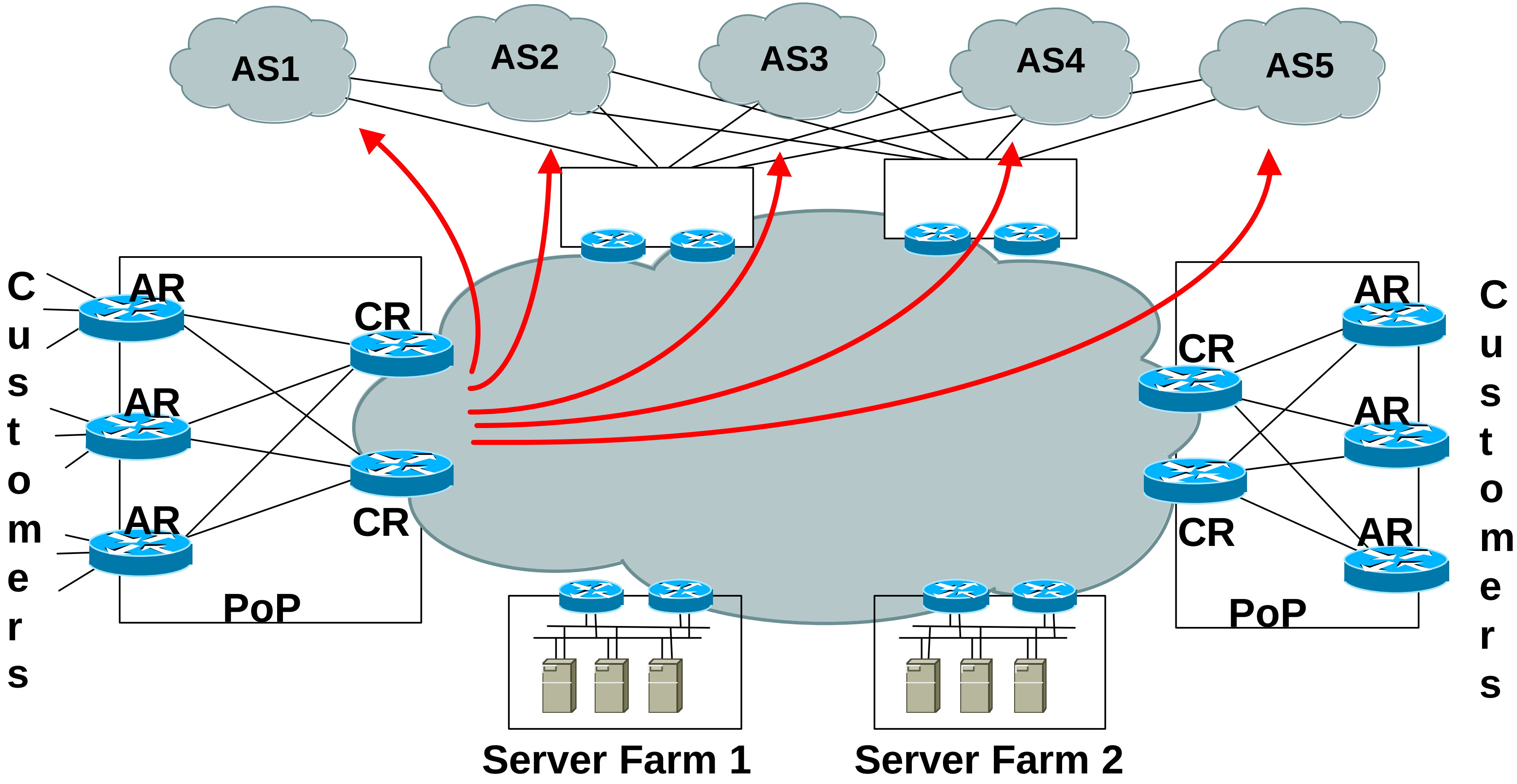
We need the Internal Traffic Matrix

Cisco.com



- “PoP to PoP”, the POP being the AR or CR

The External Traffic Matrix is a plus



- From “PoP to BGP AS”, the POP being the AR or CR
- The external traffic matrix can influence the internal one

NetFlow BGP Next Hop TOS Aggregation Issue

- Only for IP packets (IP to IP or IP to MPLS)

  Example: If a MPLS core starting from the AR, Will generate flow records from all the AR

  Note: if want to/must enable on the CR,
  investigate MPLS aware NetFlow

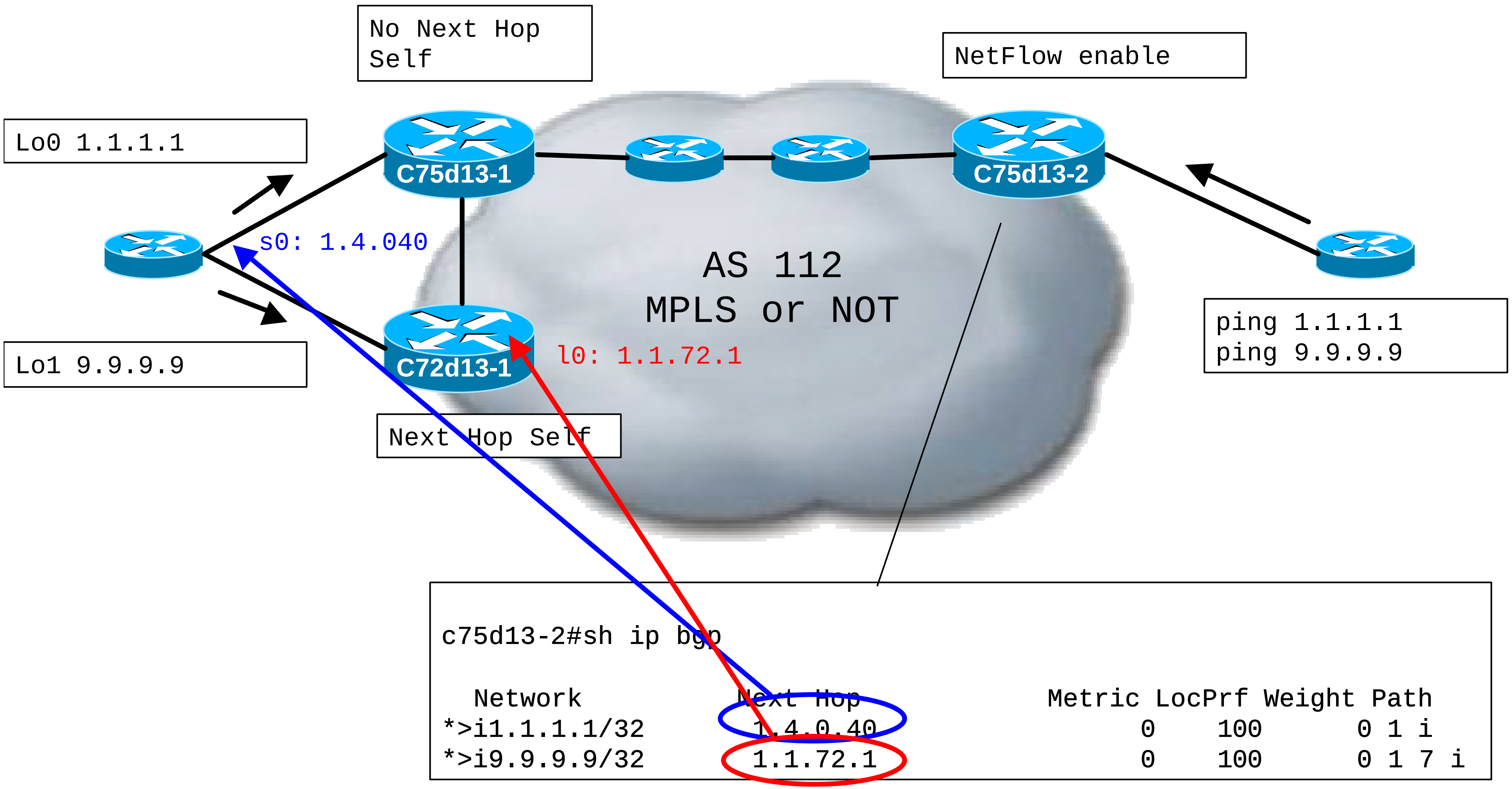
- For non BGP routes, the BGP Next Hop will be set to 0.0.0.0

  In other words, no traffic matrix for non BGP routes

NetFlow BGP Next Hop TOS Aggregation Configuration

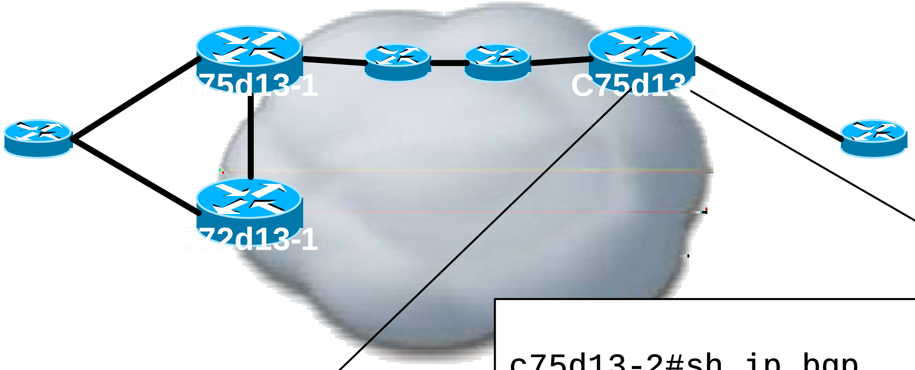
```
Router(config)#ip flow-export version 9 [origin-as | peer-as]
[bgp-nexthop]

Router(config)#ip flow-export destination <dest IP> <dest
udp-port>

Router(config)#ip flow-export source <interface>


Router (config)#ip flow-aggregation cache bgp_nexthop_tos

Router (config-flow-cache)#enabled


Router (config-if)#ip route-cache flow
```

NetFlow BGP Next Hop TOS Aggregation Testing



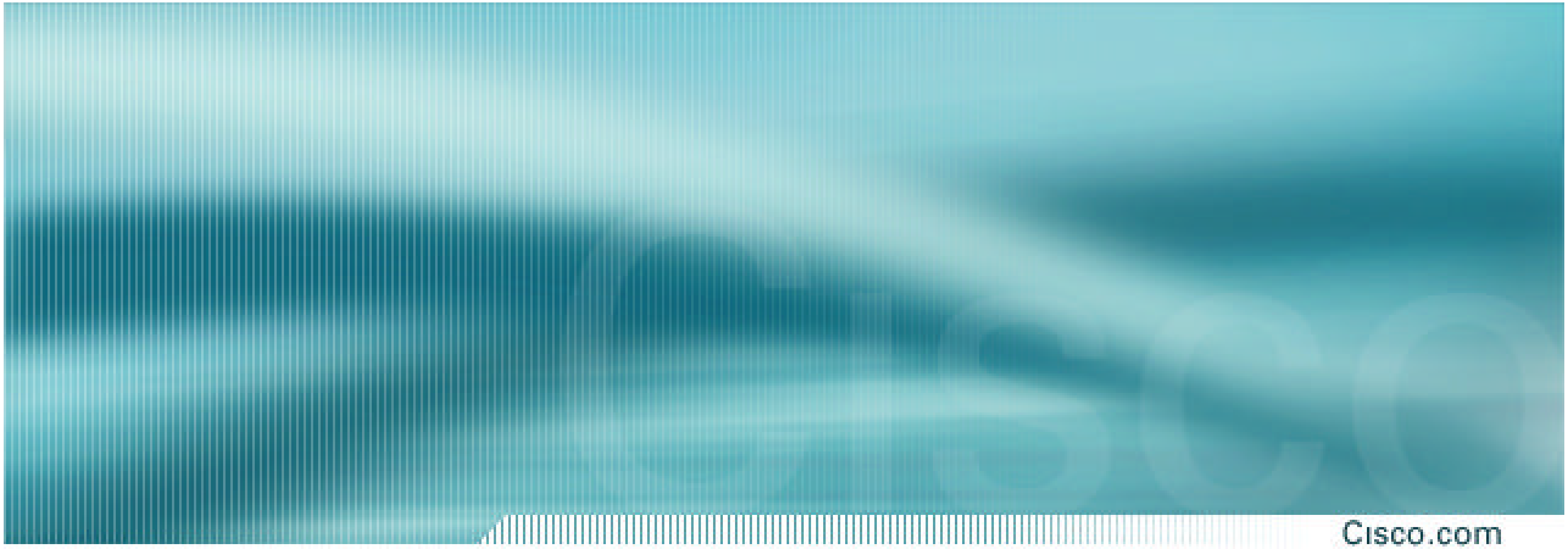
```
c75d13-2#sh ip bgp

  Network          Next Hop            Metric LocPrf Weight Path
*>i1.1.1.1/32      1.4.0.40                 0    100      0 1 i
*>i9.9.9.9/32      1.1.72.1                 0    100      0 1 7 i
```

NetFlow BGP Next Hop TOS Aggregation Testing



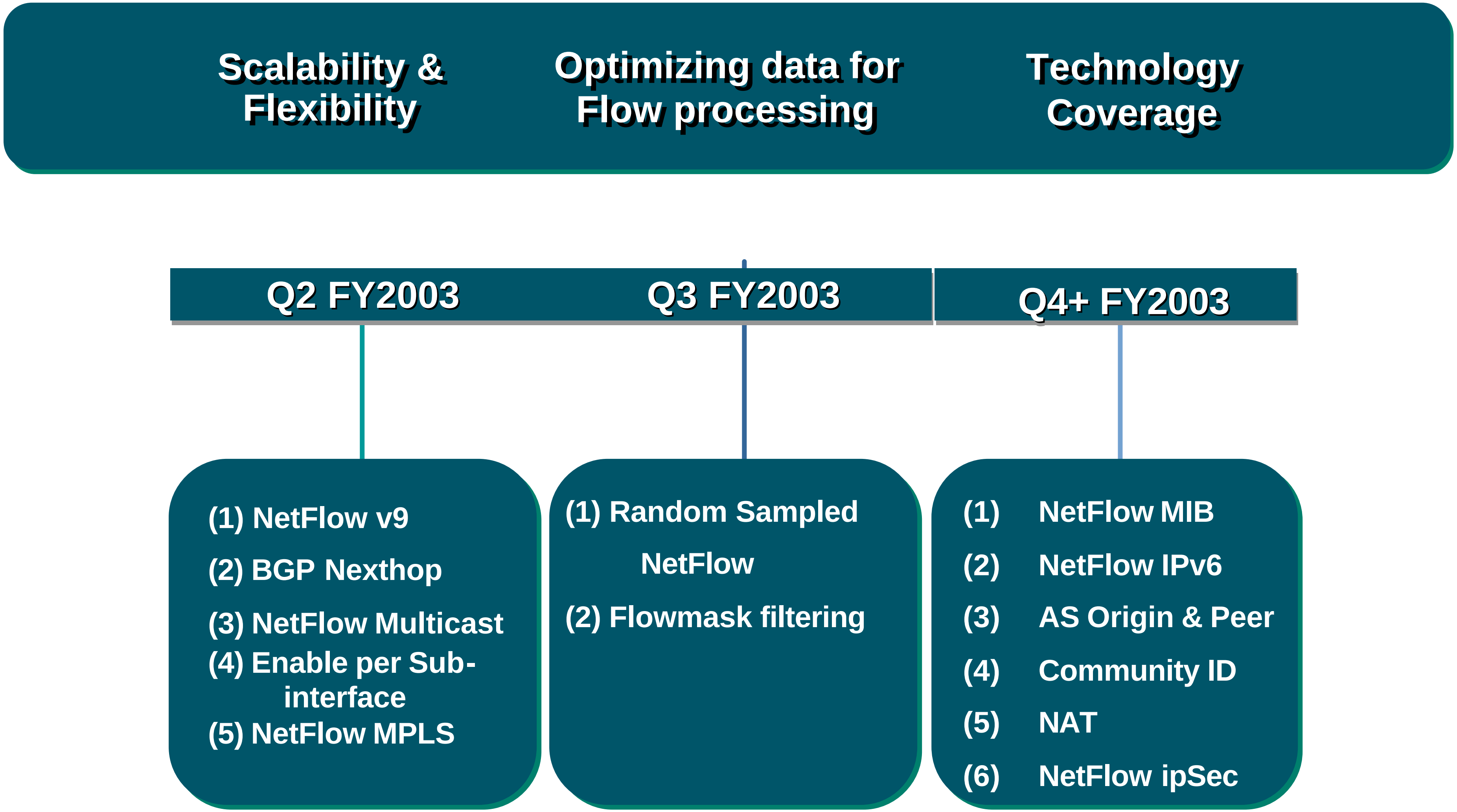
```
c75d13-2#sh ip bgp

  Network          Next Hop            Metric LocPrf Weight Path
*>i1.1.1.1/32       1.4.0.40                0    100      0 1 i
*>i9.9.9.9/32       1.1.72.1                0    100      0 1 7 i
```

```
sh ip cache verbose flow aggregation bgp-nexthop-tos

Src If         Src AS  Dst If         Dst AS  TOS Flows   Pkts  B/Pk  Active
BGP NextHop
Et1/0/1           2    Et1/0/2           1    00      1      5   100     0.0
BGP: 1.4.0.40          FOR A PING TO 1.1.1.1

Src If         Src AS  Dst If         Dst AS  TOS Flows   Pkts  B/Pk  Active
BGP NextHop
Et1/0/1           2    Et1/0/2           1    00      1      5   100     0.0
BGP: 1.1.72.1          FOR A PING TO 9.9.9.9
```



Roadmap and Future Directions

External Roadmap for NetFlow



Scalability & Flexibility

Optimizing data for Flow processing

Technology Coverage

Q2 FY2003

(1) NetFlow v9

(2) BGP Nexthop

(3) NetFlow Multicast

(4) Enable per Sub-interface

(5) NetFlow MPLS

Q3 FY2003

(1) Random Sampled NetFlow

(2) Flowmask filtering

Q4+ FY2003

(1) NetFlow MIB

(2) NetFlow IPv6

(3) AS Origin & Peer

(4) Community ID

(5) NAT

(6) NetFlow ipSec

Future Directions

- Cat6000/7600

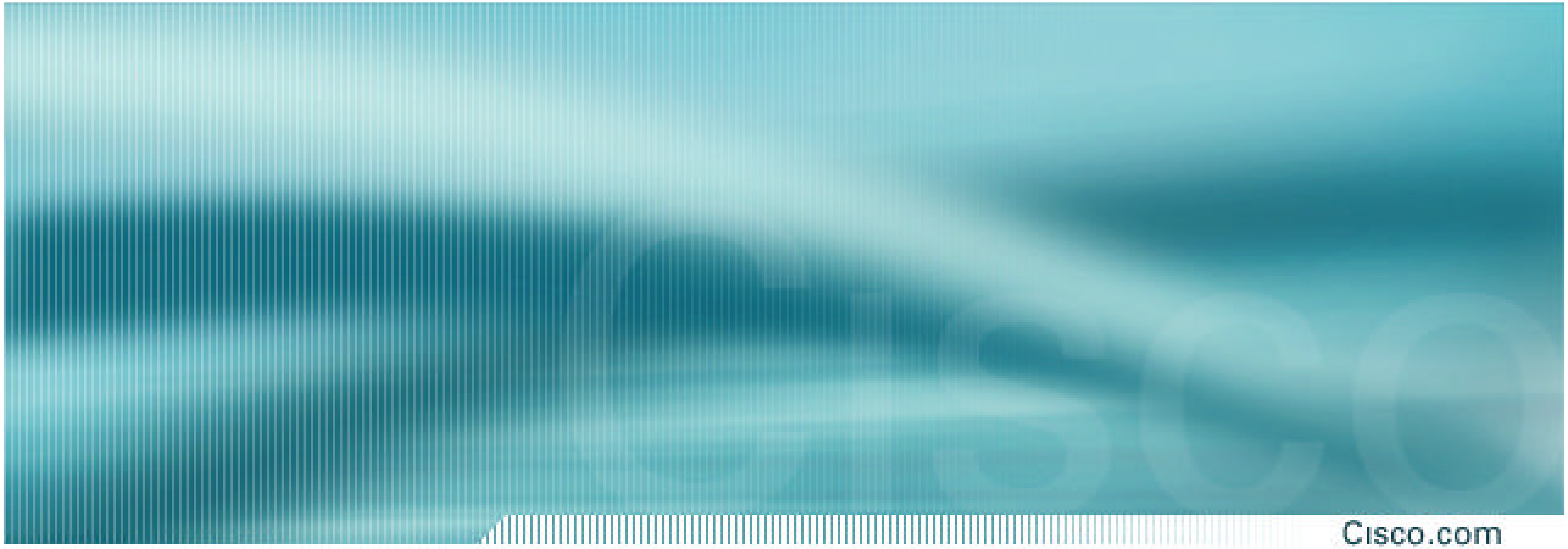
  Version 8 for the native mode

  Native mode will support dual export

  Add support for version 9

- Cat4000

  NetFlow should be supported very soon



NetFlow FlowCollector

NetFlow FlowCollector

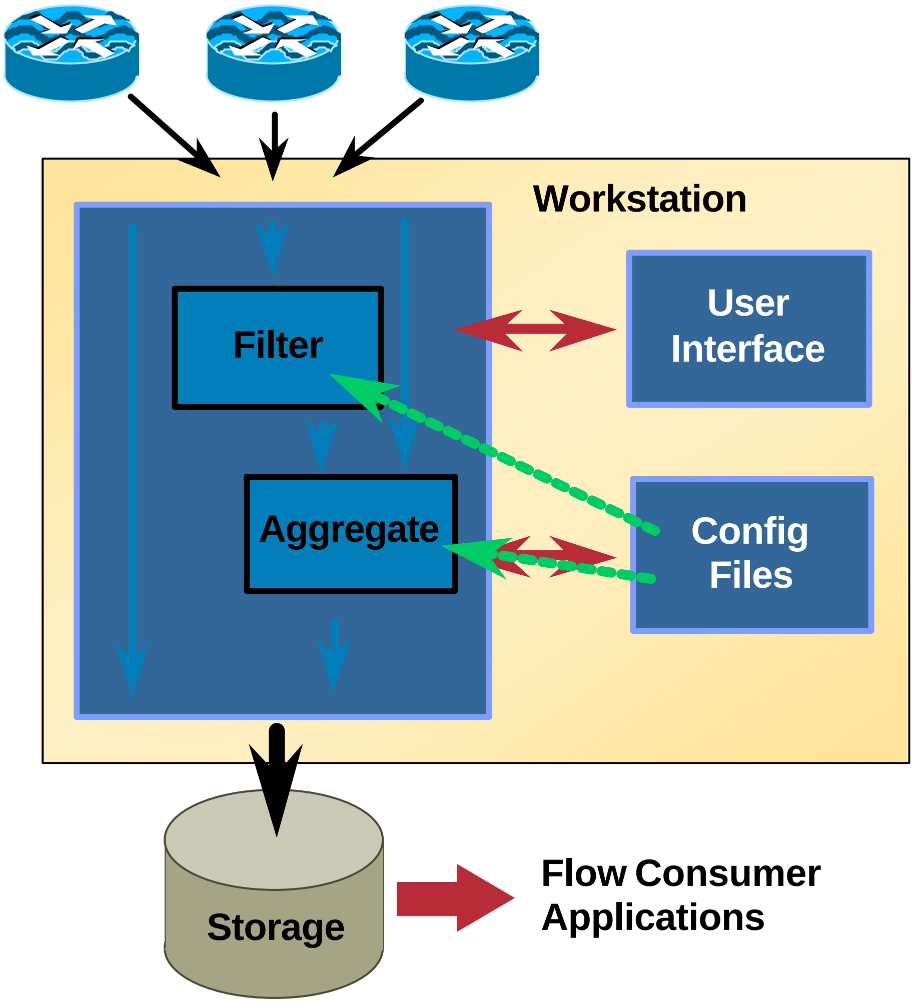
- Flow record reception
- Data volume reduction

    Filtering, Aggregation

- Flexible thread language
- File storage

    Flat or binary and compression in 3.0

- File cleanup
- Solaris and HP-UX
- No flow de-duplication

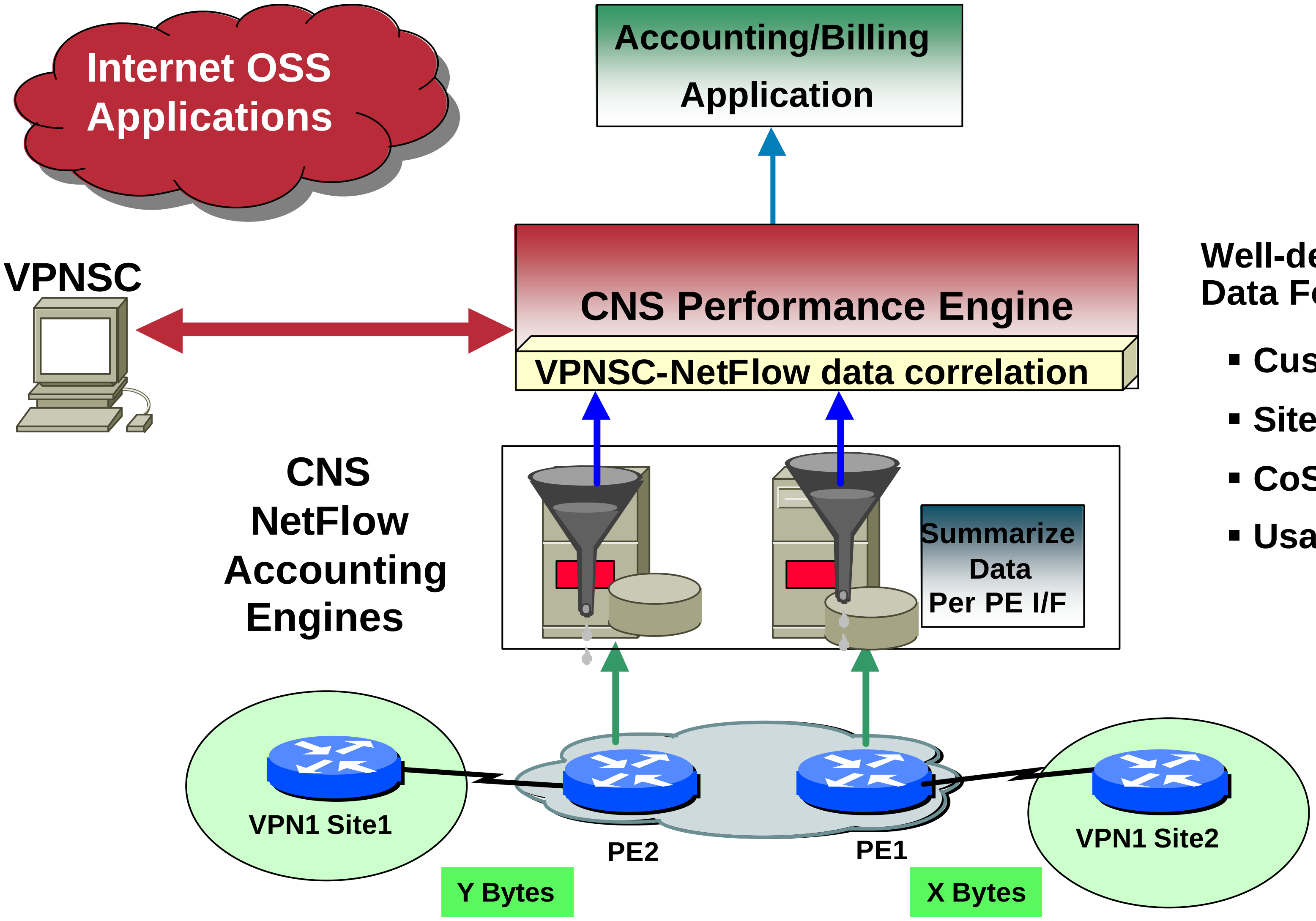


Flow Consumer Applications

New Feature in NetFlow FlowCollector 4.0

- Support NF V9 data format and templates (inc. new fields)
- Support user-configurable aggregation schemes
  - All formats v5 -> v9
- XML message set
- CNS bus support
- Deployment as Linux appliance (Redhat 7.2/IE21xx)
- Performance benchmarking document (double throughput compared to NFC 3.6)
- Already available

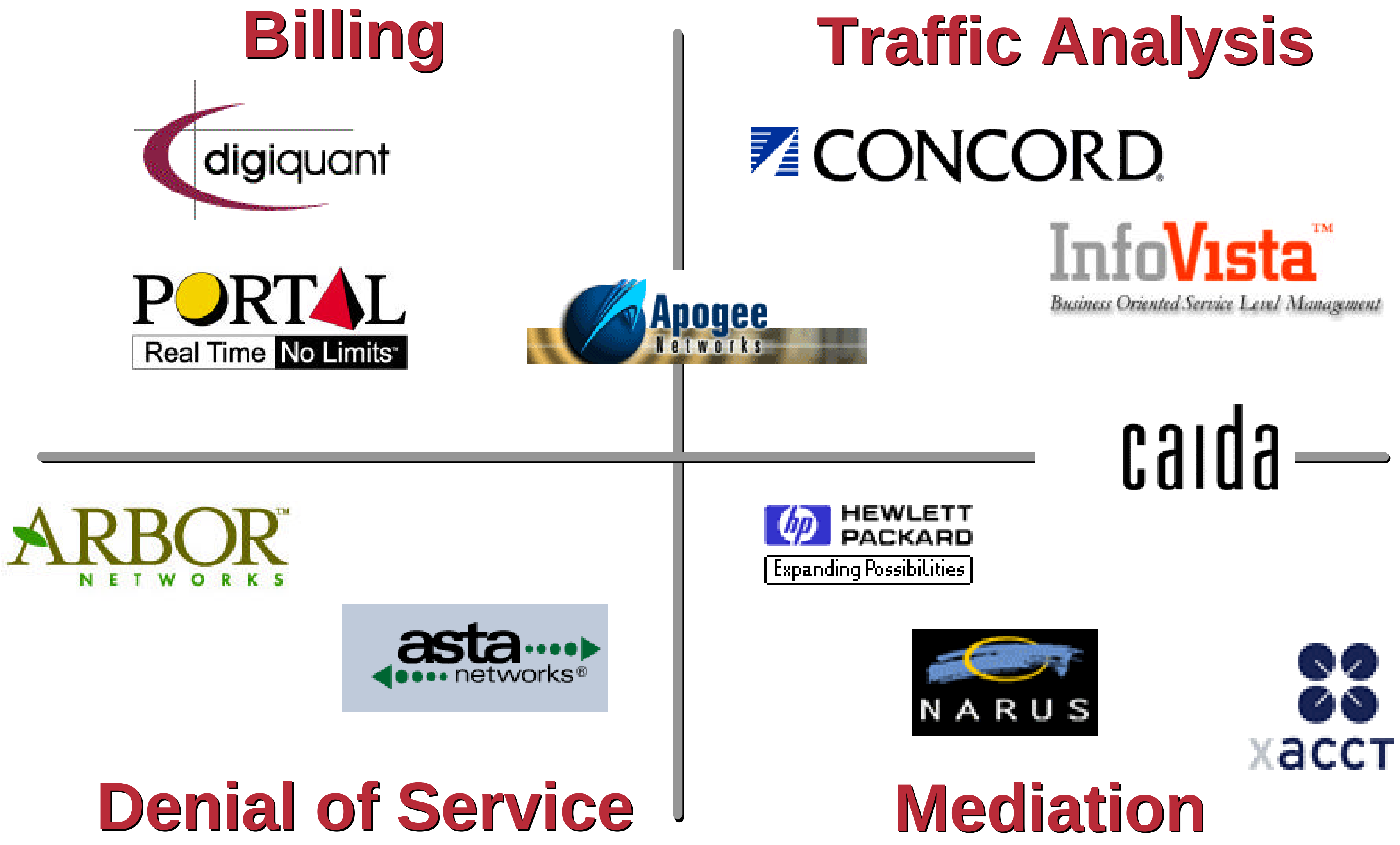
Per VPN Usage-based Accounting using CNS Performance Engine

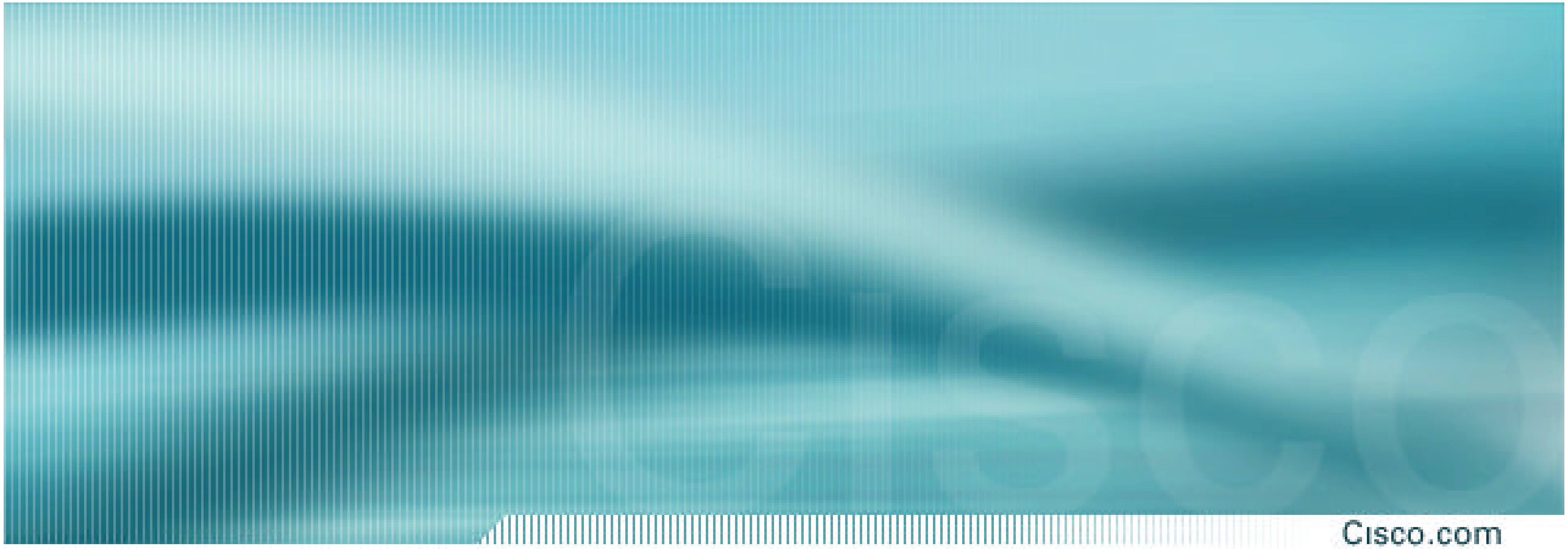

Internet OSS Applications
Accounting/Billing Application
VPNSC
CNS Performance Engine
VPNSC-NetFlow data correlation
CNS NetFlow Accounting Engines
Summarize Data Per PE I/F
VPN1 Site1
PE2
PE1
VPN1 Site2
Y Bytes
X Bytes

Well-defined Data Format

- Customer
- Site
- CoS
- Usage

NetFlow Partners

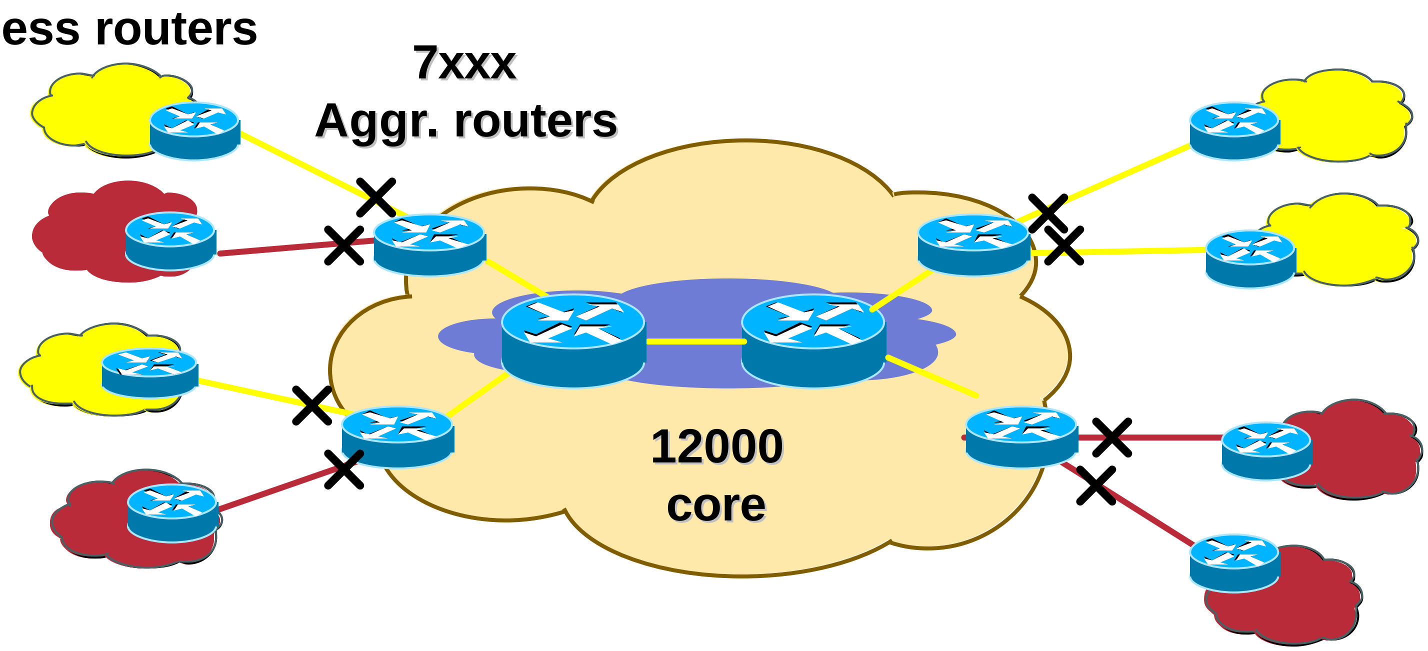




Deployment Guide

Where to deploy Netflow?

Billing

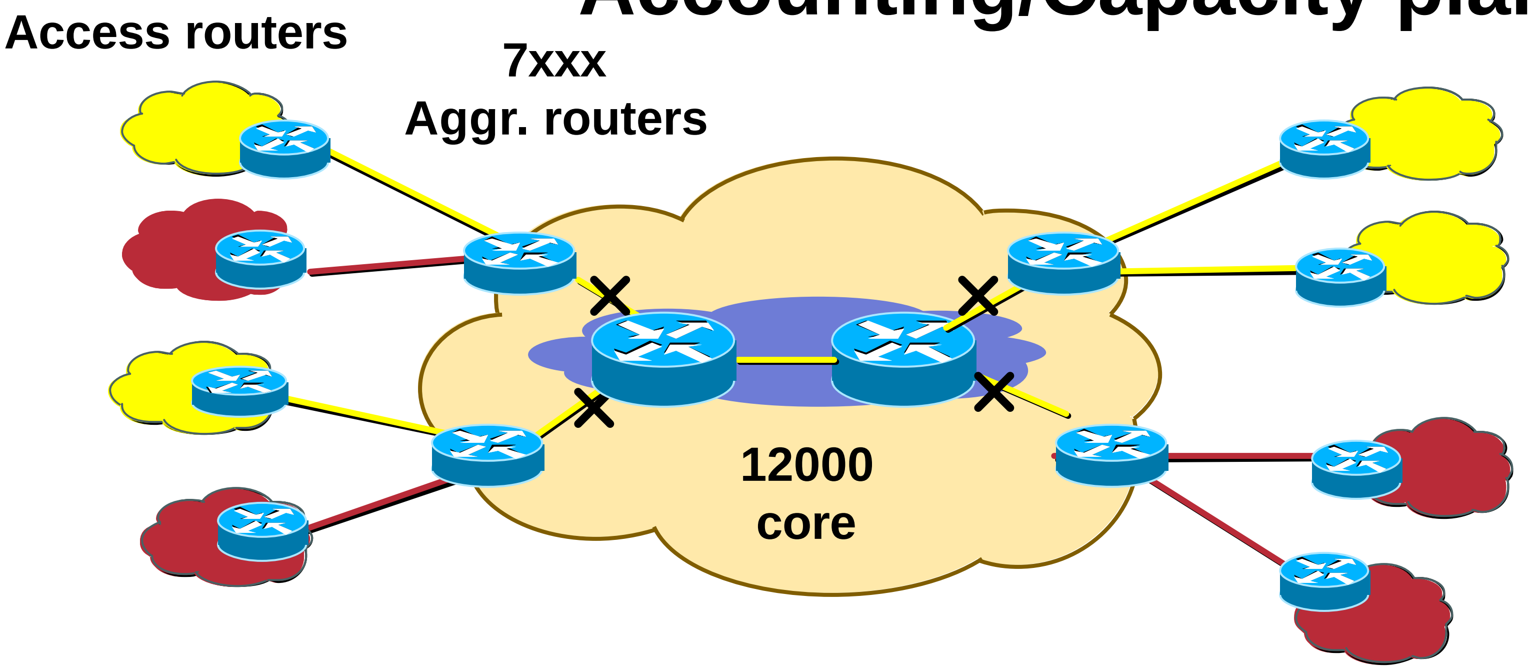


Full NetFlow

Where to deploy Netflow?

Cisco.com

Accounting/Capacity planning



Full or sampled NetFlow

Where to deploy Netflow?

- On the “edges” of the network.
- All routers because NetFlow accounts incoming traffic only
- For billing, on the aggregation routers because some GSR line cards only support sampled NetFlow.
- For accounting, capacity planning, on the aggregation routers or the GSR. Sampled NetFlow could be sufficient.

Where to deploy Netflow?

- For BGP informations, on the BGP peering routers
- Can monitor one link, egress and ingress, but should be on a MPLS PE-CE link.
- Basic principles:

  Avoid a flow duplication design. Netflow Collector doesn't do flow de-duplication. Done by partner tools

  Don't account your exported data

How many NetFlow Collector (NFC)?

- In theory, one collector per POP or Aggregation Router (7x00 router)
- For VPNSC (MPLS VPN environment), we advice one Collector per PE
- Basic principles:

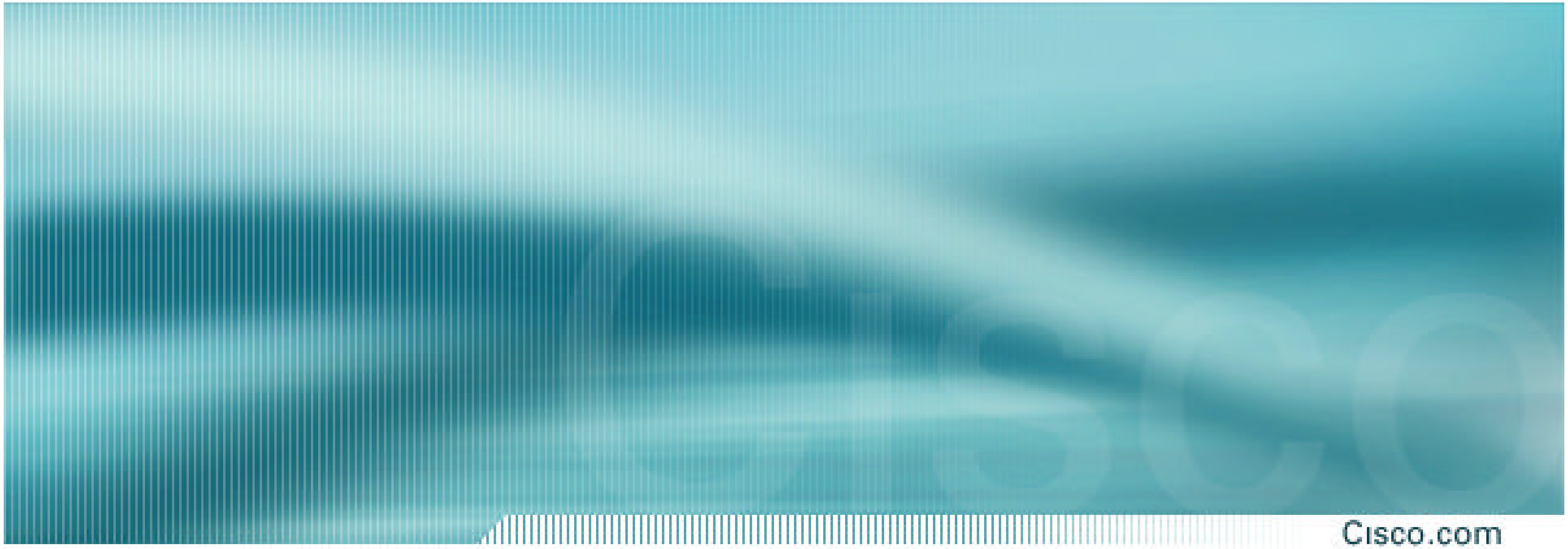
  Check your Sun capabilities

  NFC sizer calculator. Reduce the number of routers per NFC if needed.

  Rule of thumb: 10 routers per NFC

Deployment Tricks

- Enable the ifIndex persistence if accounting per interface
- Look at the router cpu (<60%) and memory before enabling NetFlow
- Check the export link bandwidth
- Use a dedicated export lan
- If you export too much traffic:
  go for the aggregations, don't export version 5
  go for sampled if on a GSR
  increase the aggregations timers
- Access-lists still account the traffic



References

NetFlow References

- **Netflow Services and Applications**

  http://www.cisco.com/go/netflow

- **Cisco Netflow Technologies Partner**

  http://www.cisco.com/warp/customer/732/partners/nfpartner.html

- **Cisco Netflow Collector/Analyzer**

  http://www.cisco.com/univercd/cc/td/doc/product/rtrmgmt/index.htm

NetFlow References

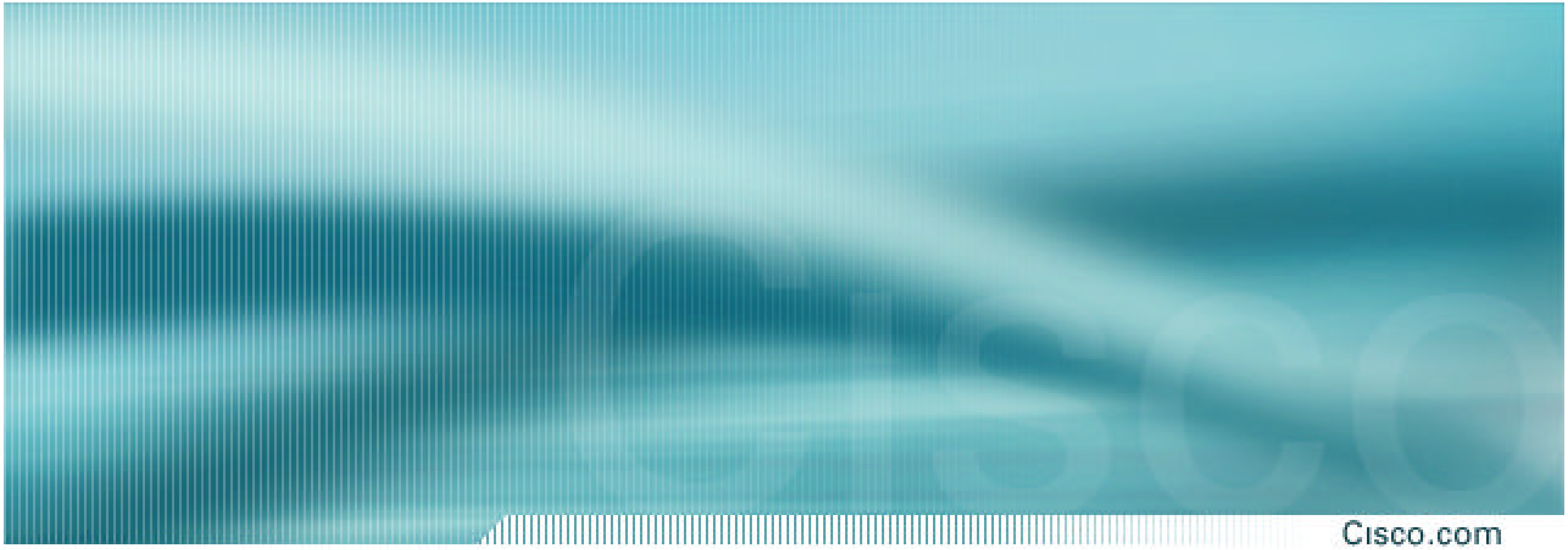
- A complete white paper

  http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/netflsol/nfwhite.htm

- An official Cisco Course (2 days)

  NetFlow Service Advanced

Questions?



NetFlow Services

Benoit Claise

bclaise@cisco.com

RIPE 44, Amsterdam

Cisco Systems ®