

# OpenFlow En Profundidad

Este documento es el resultado del trabajo del Network Startup Resource Center (NSRC en <http://www.nsrc.org>) y del Indiana Center for Network Translational Research and Education (InCNTRE). Este documento puede ser libremente copiado, modificado y reutilizado con la condición de que cualquier reutilización debe reconocer al NSRC y al InCNTRE como las fuentes originales.



UNIVERSITY OF OREGON



- Estas diapositivas tienen contenido significativo distribuidos por:



# Estado actual de los Estándares

- OpenFlow 1.3
  - Apoyado y Mantenido
  - Ahora corresponde a 1.3.5
- OpenFlow 1.5.1 últimas especificaciones liberadas
- OF-Config 1.2 Disponible
- Pruebas de Conformidad
  - 1.0.1 Completo, L2, L3 Perfiles
  - 1.3.4 Tabla de perfil básico
  - <https://www.opennetworking.org/openflow-conformance-certified-products>



# OpenFlow 1.0 a 1.3



UNIVERSITY OF OREGON



# Más Tablas de Flujo

## 1.0

- Tabla de Flujo Individual

## 1.3

- Permite múltiples Tablas de Flujo
  - Saltar a otras tablas, pero sólo en dirección positiva.
- Incluye una tabla de grupos con tipos de tabla de grupo múltiple
- Incluye medidor de tabla para la limitación de velocidad y calidad de servicio (QoS)



# Tabla de Flujo - Tipos

1.0

Match Fields	Priority	Counters	Actions	Timeouts	Cookies
--------------	----------	----------	---------	----------	---------

1.3

Match Fields	Priority	Counters	Instructions	Timeouts	Cookies
--------------	----------	----------	--------------	----------	---------

Group Tables

Group ID	Type	Counters	Action Buckets
----------	------	----------	----------------

Meter Tables

Meter Identifier	Meter Bands	Counters
------------------	-------------	----------



# Procesamiento de paquetes

1.0 - Si el paquete coincide con la entrada de la tabla de flujo, realice “Acción” .....

1.3 - Si el paquete coincide con la entrada de la tabla de flujo, mira “Instrucciones” .....

- Las instrucciones pueden ejecutar acciones inmediatas (s), o
- Las instrucciones pueden establecer acciones en el conjunto de acciones.
- Las instrucciones pueden escribir en el campo de metadatos
- Las instrucciones pueden cambiar procesamiento en las tuberías:
  - Vaya (Goto) a tabla X
  - Vaya (Goto) a la tabla de grupo Y
  - Aplica el medidor Z

Nueva Estructura de datos de la tubería



# Instrucciones

Instruccion	Descripcion	Obligatoria / Opcional
OFPIT_GOTO_TABLE	Saltar a la tabla n	Obligatoria
OFPIT_WRITE_METADATA	Escribir en el campo de metadato para usar luego en la tuberia.	Opcional
OFPIT_WRITE_ACTIONS	Agregar acciones al conjunto de acciones para su uso posterior	Obligatoria
OFPIT_APPLY_ACTIONS	Aplicar acciones inmediatamente	Opcional
OFPIT_CLEAR_ACTIONS	Borra todas las acciones del conjunto de acciones	Opcional
OFPIT_METER	Aplicar medidor (Limitador de velocidad)	Opcional
OFPIT_EXPERIMENTER	(0xFFFF) Instrucciones experimentales	Opcional



# Campos requeridos (propuesto)

Match Field Campo	Descripcion	Pre-requisito
OXM_OF_IN_PORT	Puerto de entrada	Ninguno
OXM_OF_ETH_DST	Dirección Ethernet de destino	Ninguno
OXM_OF_ETH_SRC	Dirección Ethernet de origen	Ninguno
OXM_OF_ETH_TYPE	Tipo de Ethernet de la carga útil del paquete OpenFlow	Ninguno
OXM_OF_IP_PROTO	IPv4 o IPv6 número de protocolo	ETH_TYPE = 0x800 or ETH_TYPE = 0x86dd
OXM_OF_IPV4_SRC	Dirección de origen IPv4	ETH_TYPE = 0x800
OXM_OF_IPV4_DST	Dirección de destino IPv4	ETH_TYPE = 0x800
OXM_OF_IPV6_SRC	Dirección de origen IPv6	ETH_TYPE = 0x86dd
OXM_OF_IPV6_DST	Dirección de destino IPv6	ETH_TYPE = 0x86dd
OXM_OF_TCP_SRC	Puerto de origen TCP	IP_PROTO = 6
OXM_OF_TCP_DST	Puerto de destino TCP	IP_PROTO = 6
OXM_OF_UDP_SRC	Puerto de origen UDP	IP_PROTO = 17
OXM_OF_UDP_DST	Puerto de destino UDP	IP_PROTO = 17

# Campos opcionales (propuesto)

Match Field	Description	Match Field	Description
OXM_OF_IN_PHY_PORT	Puerto de entrada física	OXM_OF_ARP_THA	Target L2 Add of ARP payload
OXM_OF_METADATA	Tabla de metadata	OXM_OF_IPV6_FLABEL	IPv6 flow label
OXM_OF_VLAN_VID	VLAN ID	OXM_OF_ICMPV6_TYPE	ICMPv6 type
OXM_OF_VLAN_PCP	VLAN PCP	OXM_OF_ICMPV6_CODE	ICMPv6 code
OXM_OF_IP_DSCP	Diff Serv Code Point	OXM_OF_IPV6_ND_TARGET	Target Addr IPv6 Neigh Disc
OXM_OF_IP_ECN	bits Ecn de la cabecera IP	OXM_OF_IPV6_ND_SLL	SRC L2 Addr IPv6 Neigh Disc
OXM_OF_SCTP_SRC	SCTP puerto de origen	OXM_OF_IPV6_ND_TLL	Target L2 Add IPv6 Neigh Disc
OXM_OF_SCTP_DST	SCTP puerto de destino	OXM_OF_MPLS_LABEL	MPLS label
OXM_OF_ICMPV4_TYPE	Tipo de ICMP	OXM_OF_MPLS_TC	MPLS TC
OXM_OF_ICMPV4_CODE	Codigo ICMP	OXM_OF_MPLS_BOS	MPLS BoS bit
OXM_OF_ARP_OP	ARP opcode	OXM_OF_PBB_ISID	I-SID in PBB svc instance tag
OXM_OF_ARP_SPA	Direccion IPv4 SRC de la carga util ARP	OXM_OF_TUNNEL_ID	Metadata for a logical port
OXM_OF_ARP_TPA	Target IPv4 Addr of ARP payload	OXM_OF_IPV6_EXTHDR	IPv6 extension header field
OXM_OF_ARP_SHA	SRC Eth Addr of ARP payload		

# 1.3 Acciones

Action	Description
OFPAT_OUTPUT	Output to port
- PORT #	Physical or Logical switch port #
- ALL, CONTROLLER, TABLE, IN_PORT	Required Reserved Ports
- LOCAL, NORMAL, FLOOD	Optional Reserved ports
OFPAT_COPY_TTL_OUT	Copy TTL to next outwards header
OFPAT_COPY_TTL_IN	Copy TTL to next inwards header
OFPAT_SET_MPLS_TTL	Set MPLS TTL
OFPAT_DEC_MPLS_TTL	Decrement MPLS TTL
OFPAT_PUSH_VLAN	Push a new VLAN Tag
OFPAT_POP_VLAN	Pop the outer VLAN Tag
OFPAT_PUSH_MPLS	Push a new MPLS Tag
OFPAT_POP_MPLS	Pop the outer MPLS Tag

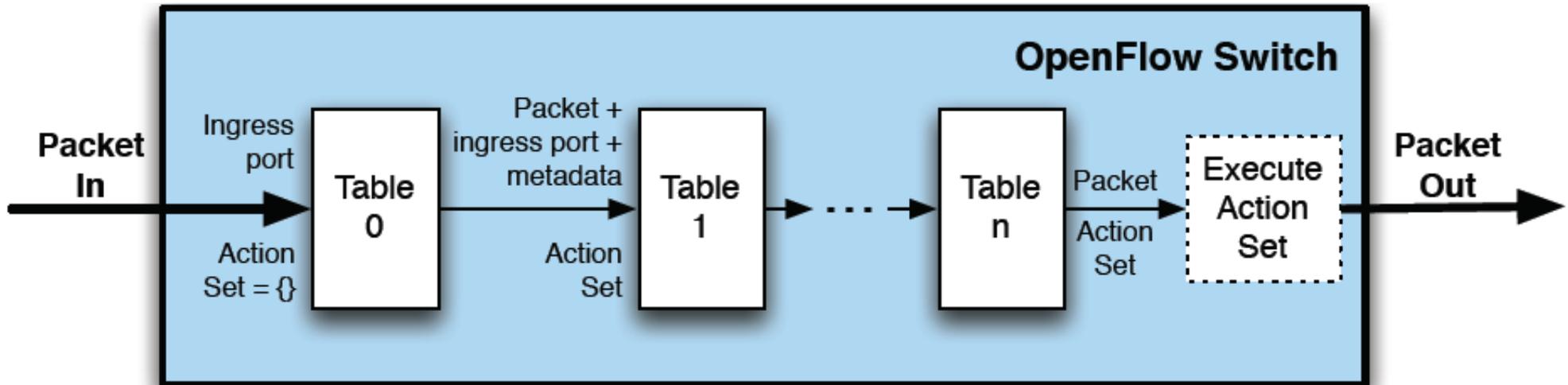


# Mas acciones de 1.3

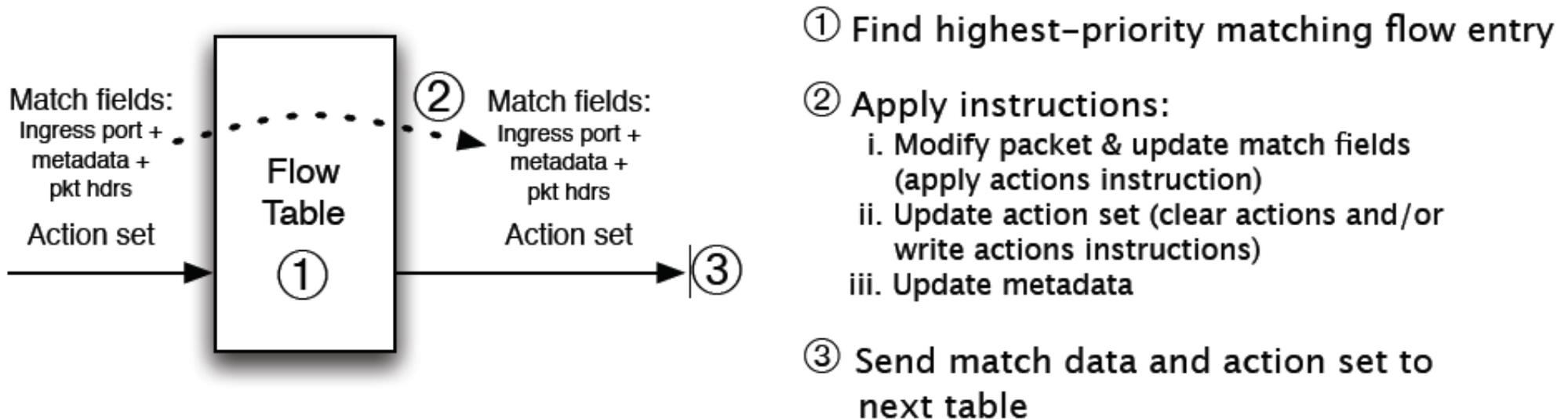
Accion	Descripcion
OFPAT_SET_QUEUE	Set queue id when outputting to a port
OFPAT_GROUP	Apply Group
OFPAT_SET_NW_TTL	Set IP TTL
OFPAT_DEC_NW_TTL	Decrement IP TTL
OFPAT_SET_FIELD	Set a header field using OXM TLV format
OFPAT_PUSH_PBB	Push a new PBB service tag (I-TAG)
OFPAT_POP_PBB	Pop the outer PBB service tag (I-TAG)
OFPAT_EXPERIMENTER	0xffff
OFPAT_DROP	No explicit drop action



# 1.3 Packet flow through Processing Pipeline



(a) Packets are matched against multiple tables in the pipeline



(b) Per-table packet processing

# 1.3 Tipos de mensajes

- Symmetric
  - Sent without solicitation in either direction
- Controller-to-Switch
  - initiated by the controller
  - may or may not require a response from the switch.
- Asynchronous
  - sent without a controller soliciting them from a switch
  - denote a packet arrival, switch state change, or error

# 1.3 Simetrico

- OFPT\_HELLO
- OFPT\_ERROR
- OFPT\_ECHO\_REQUEST
- OFPT\_ECHO\_REPLY
- OFPT\_EXPERIMENTER

# 1.3 Controlador-al-switch

- OFPT\_FEATURES\_REQUEST
- OFPT\_FEATURES\_REPLY
- OFPT\_GET\_CONFIG\_REQUEST
- OFPT\_GET\_CONFIG\_REPLY
- OFPT\_SET\_CONFIG
- OFPT\_PACKET\_OUT
- OFPT\_FLOW\_MOD
- OFPT\_GROUP\_MOD
- OFPT\_PORT\_MOD
- OFPT\_TABLE\_MOD

# 1.3 Controlador-al-switch Mensajes (Cont.)

- OFPT\_MULTIPART\_REQUEST
- OFPT\_MULTIPART\_REPLY
- OFPT\_BARRIER\_REQUEST
- OFPT\_BARRIER\_REPLY
- OFPT\_ROLE\_REQUEST
- OFPT\_ROLE\_REPLY



# 1.3 Asíncrono

- OFPT\_PACKET\_IN
- OFPT\_FLOW\_REMOVED
- OFPT\_PORT\_STATUS

# Tabla-perdida

- Especifica cómo procesar el paquete que no coincide con ningún flujo en la tabla de flujo.
- De forma predeterminada, los paquetes sin coincidencia “match” se descartan.
- El Controlador puede modificar este comportamiento. Se identifica por coincidencia “match” y prioridad.
- Una entrada de flujo con toda la coincidencia de la mascara wildcard y prioridad 0 es una entrada de flujo de tabla-perdida.
- Debe al menos soportar el envío de paquetes al controlador.

# 1.3 Mensaje de Error

- OFPT\_ERROR utilizado para notificar al controlador de un problema.
- Cada mensaje de error tiene un tipo, el código y los datos.
- “Type” indica tipo de error de alto nivel .
- “Code” se interpreta basadas en el tipo y señala el tipo específico de problema.
- El campo de datos contiene al menos los primeros 64 bytes de la solicitud fallida que provocó el error.

# 1.3 Errores - Ejemplos

- Solicitud: `ofp_flow_mod` con puerto de salida invalido en acción.
  - Tipo de Error: `OFPET_BAD_ACTION`.
  - Codigo de Error: `OFPBAC_BAD_OUT_PORT`.
- Solicitud: `ofp_flow_mod` con “match” invalido.
  - Tipo de Error: `OFPET_BAD_MATCH`.
  - Codigo de Error: `OFPBMC_BAD_TYPE`.
- Solicitud: “Match” en `ofp_flow_mod` con pre-requisitos que faltan.
  - Tipo de Error: `OFPET_BAD_MATCH`.
  - Codigo de Error: `OFPBMC_BAD_PREREQ`.



# 1.3 Tipos de Tablas de Grupo

**todo** - Ejecutar cada cubo

- cada cubo obtiene copia del paquete
- inundaciones, multidifusión, etc.

**seleccionar** - ejecutar uno de los cubos en el grupo

- Basado en el algoritmo definido en el switch (por ejemplo, RR)
- distribución de la carga a través de un tramo de enlaces de puertos miembros.

**indirecta** - ejecutar un solo cubo en el grupo

- utilizado para los próximos saltos y convergencia rápida

**rápida conmutación por error-** ejecutar primer cubo en vivo  
(atada a la interfaz)

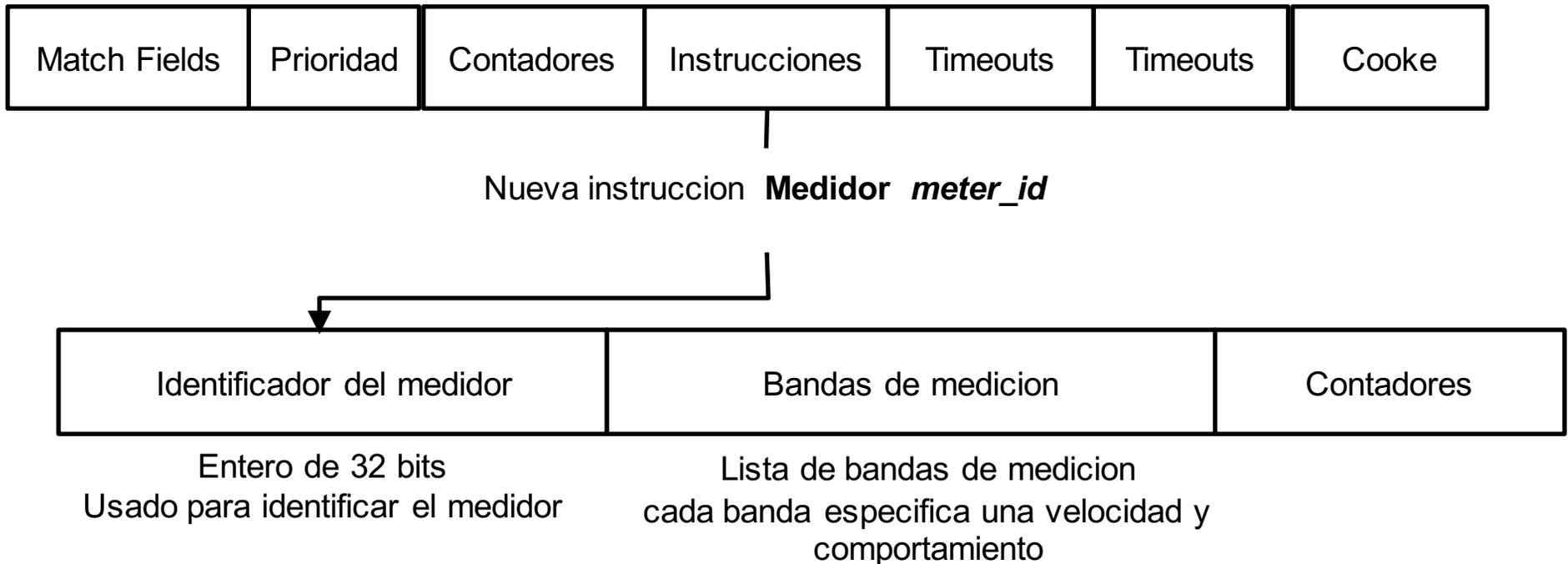
# OpenFlow Calidad de servicio QoS

## OF 1.0

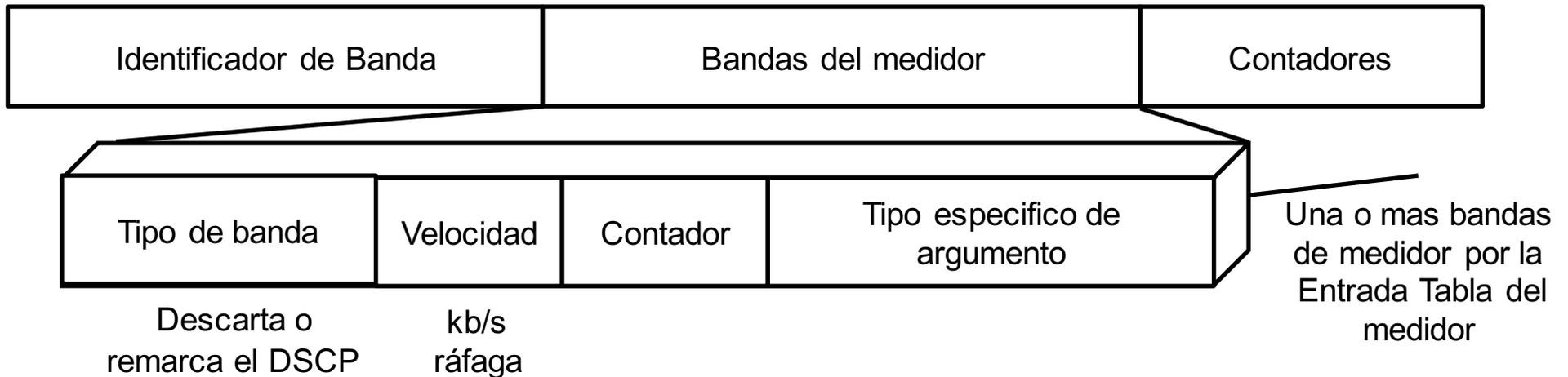
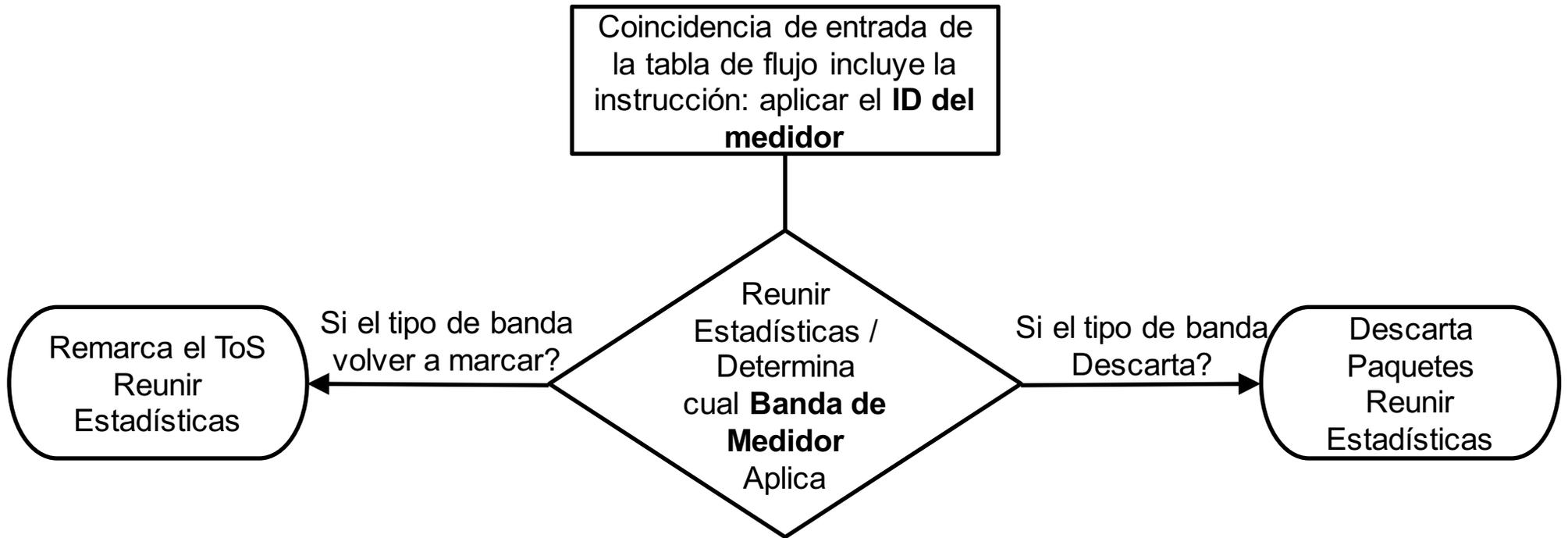
- Acción opcional "Poner en cola" paquetes de reenvío a través de una cola conectada a un puerto.
- Los campos de cabecera pueden incluir prioridad de VLAN e IP ToS, para que puedan ser emparejado en contra y re-escritos.

## OF 1.3

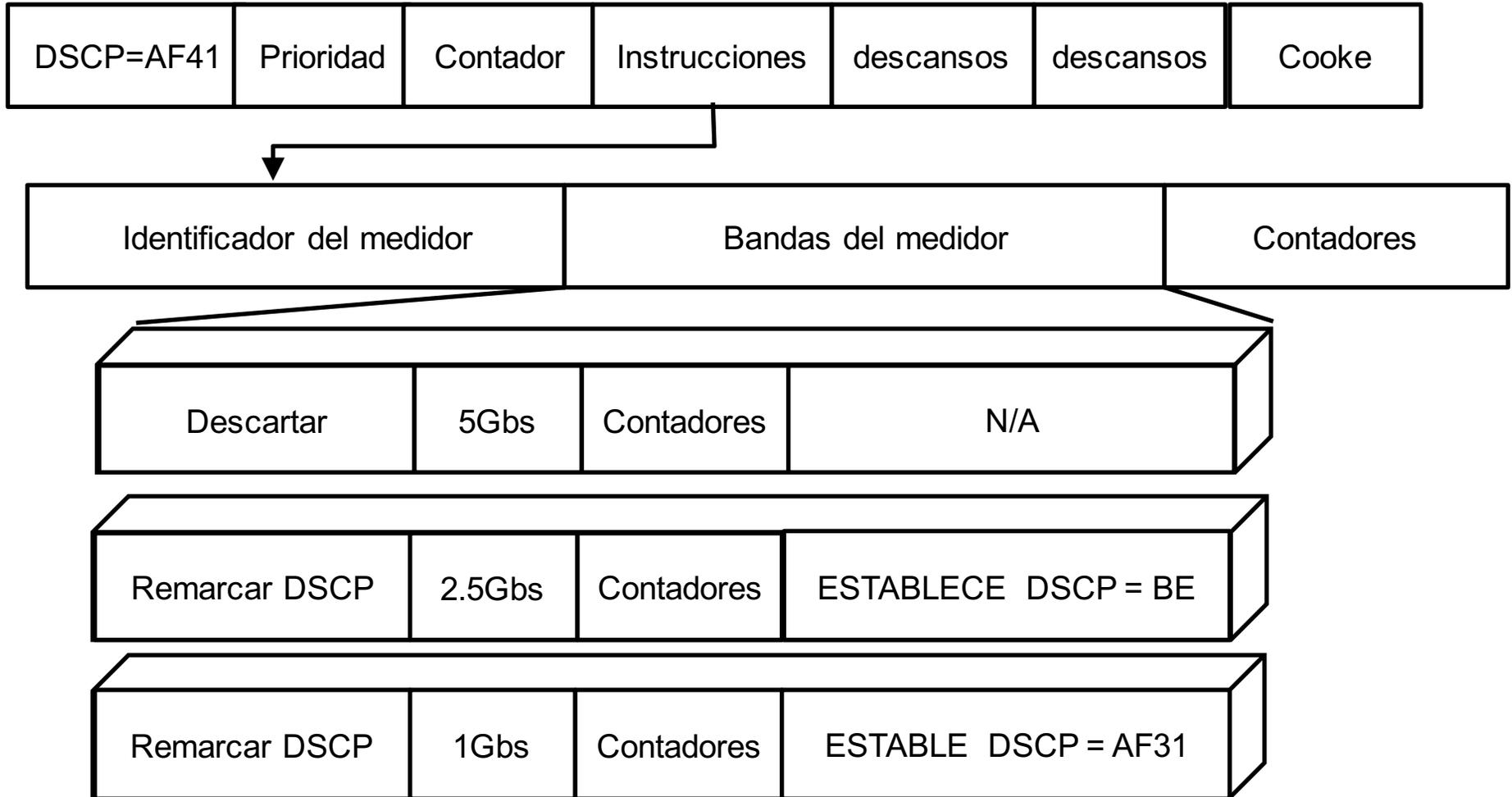
- Cosas de 1.0
- Nueva tabla "Tabla de medidor"



# 1.3 QoS (Cont.)



# 1.3 QoS (Cont.)



*" el medidor se aplica la banda de medidor con la velocidad más alta configurada y que es inferior a la velocidad de medición actual "*

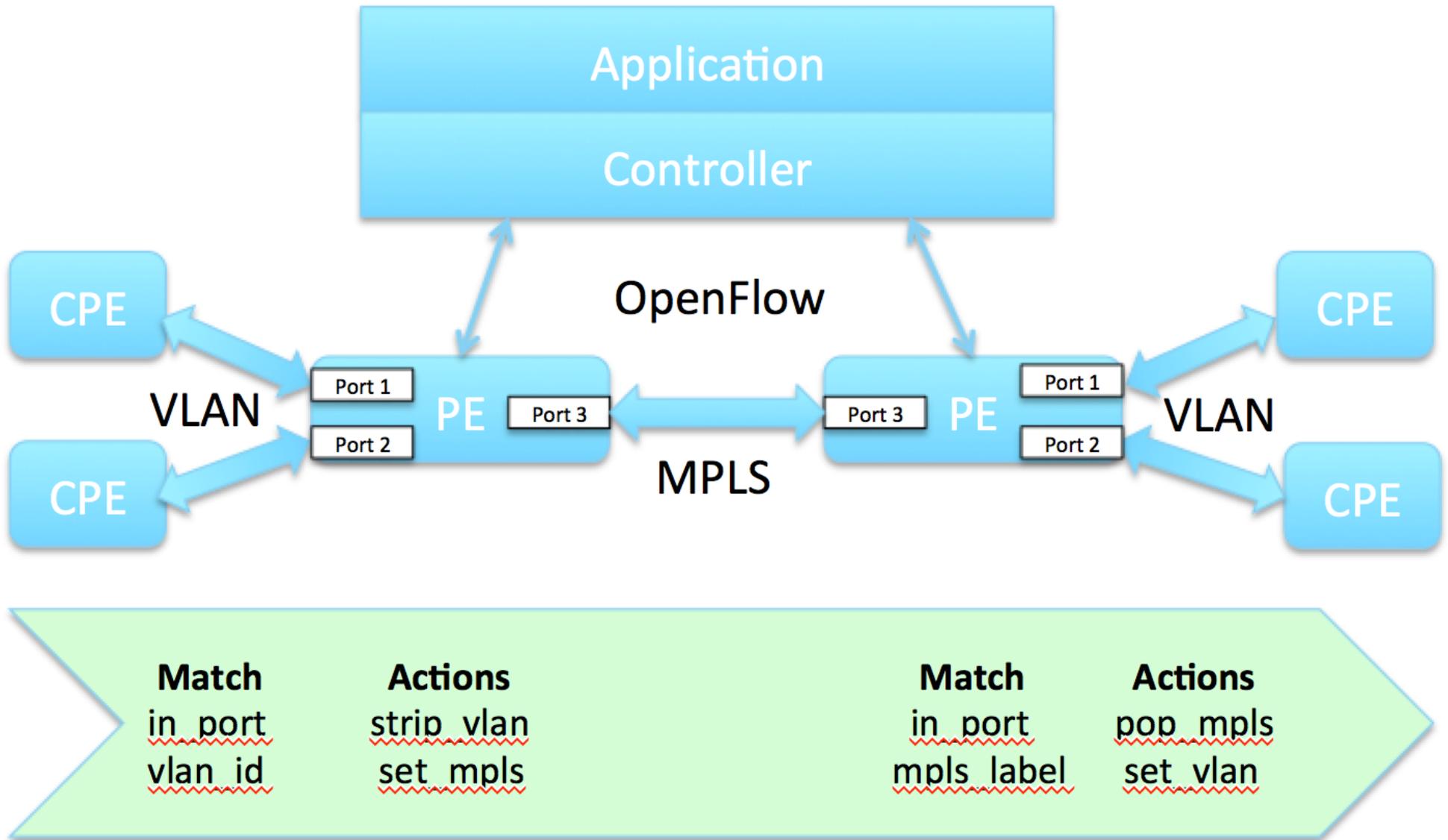
# OpenFlow 1.3

## Ejemplos y casos de uso

# 1.3 Comportamientos de red - Adicionales

- Tuneles
  - PBB/Mac en Mac
  - Paquetes Encap/Decap
- QinQ VLAN apilamiento (Stacking)
- Empujar/Pop/Re-escribir MPLS
- Emulación de enrutamiento (decremento TTL)
- Multicast/Broadcast (Tabla de Grupo)
- Funciones de experimentador???

# VLAN a MPLS Ejemplo de conversion



# La redirección HTTP usando Extensión - Ejemplo

- El controlador actúa como redirector al servidor designado
- Match
  - ip\_prot = 6
  - tcp\_dst = 80
  - extensión ***Http\_method*** para que coincida con el http "GET"
- Acción
  - Modifica ipv4\_dst del paquete
  - Envia puertos del switch hacia el servidor designado

# Ejemplo de un simple Router



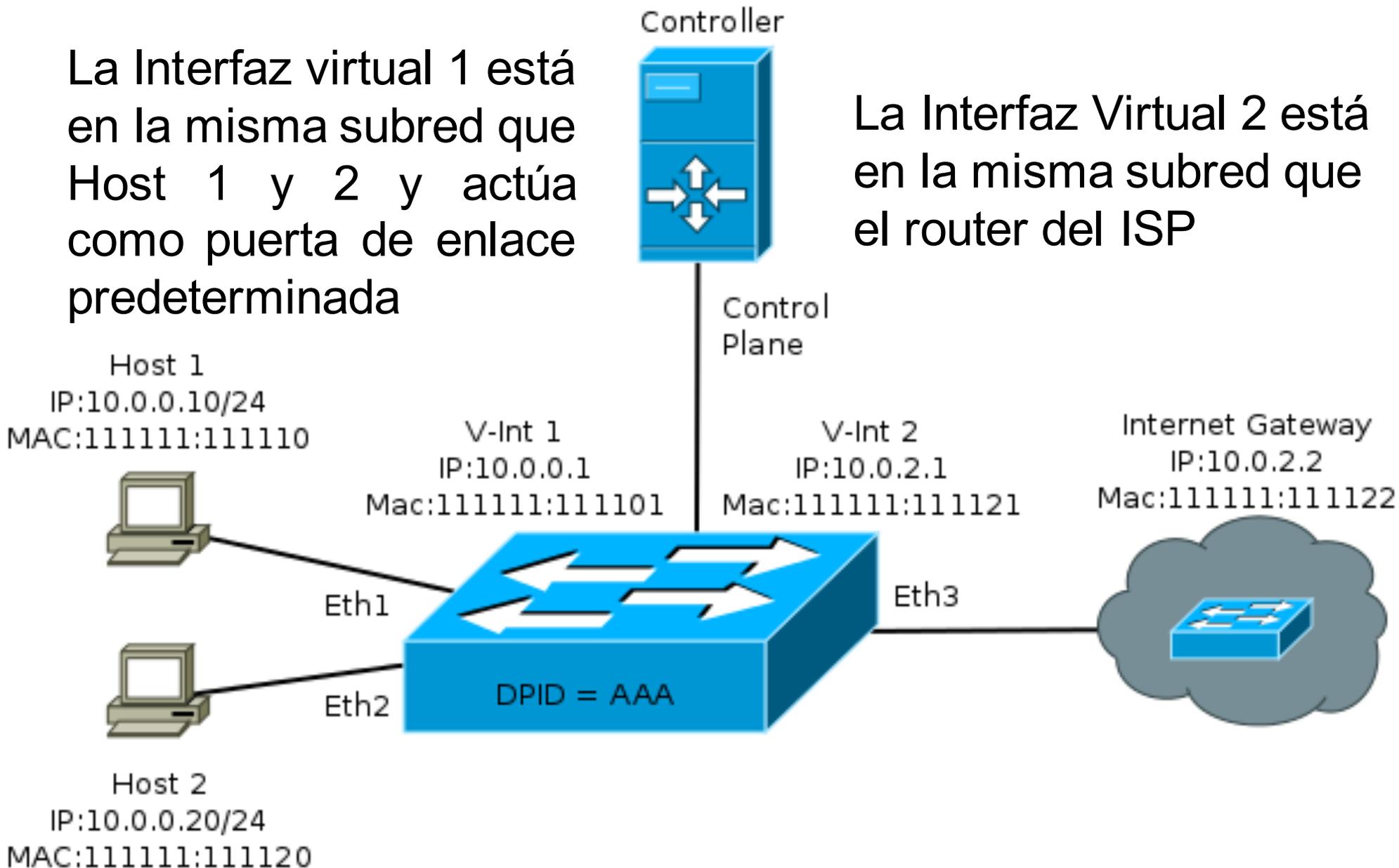
UNIVERSITY OF OREGON



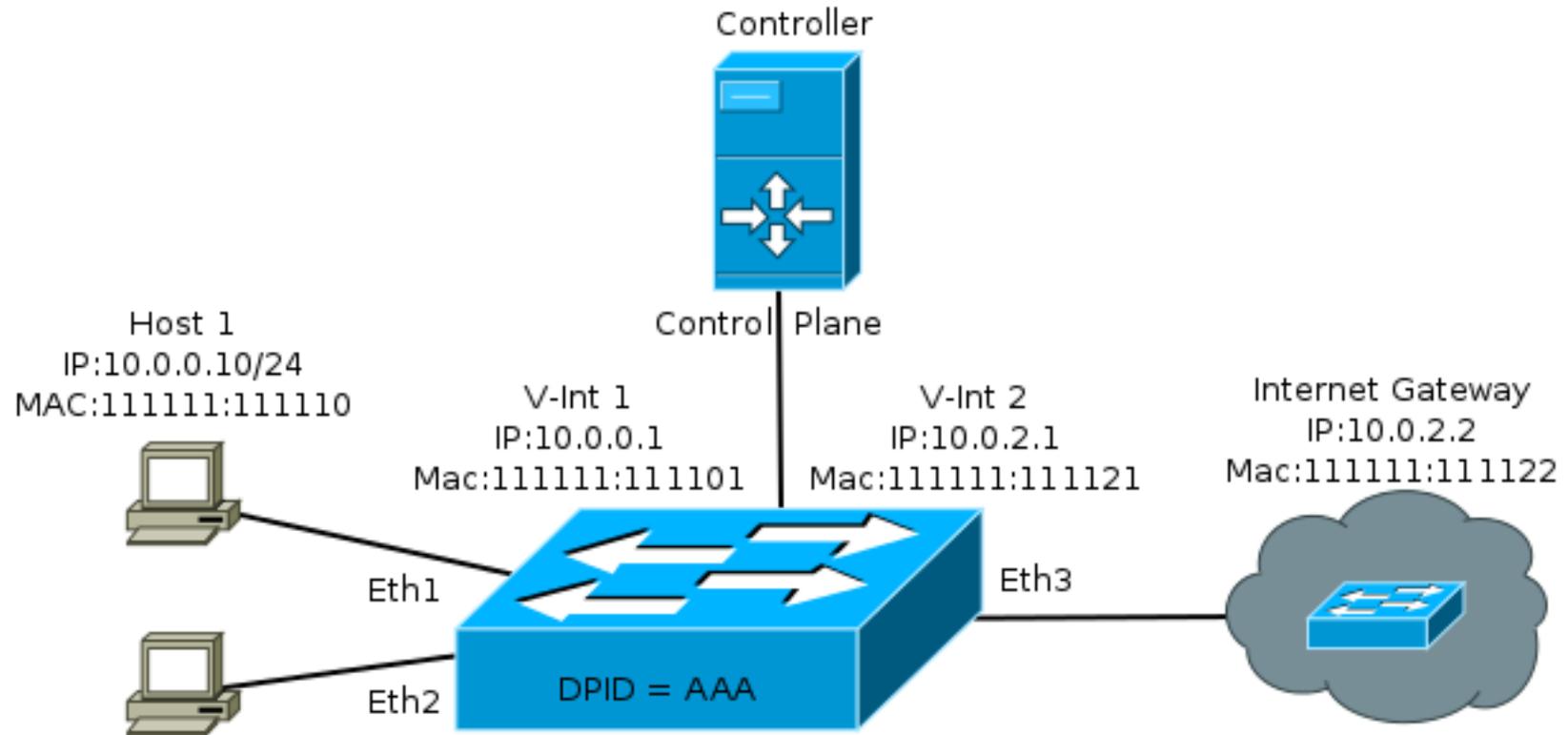
# Controlador corriendo una aplicacion del Router Virtual

La Interfaz virtual 1 está en la misma subred que Host 1 y 2 y actúa como puerta de enlace predeterminada

La Interfaz Virtual 2 está en la misma subred que el router del ISP



# Table Arp en el controlador



Arp Table in the Controller Virtual Router					
No.	IP	MAC	Switch DPID	Port	Action
1	10.0.0.1	111111:111101	-	V-Int1	
2	10.0.2.1	111111:111121	-	V-Int2	OFPT-PACKET-OUT
3	10.0.0.10	111111:111110	AAA	Eth1	OFPT-PACKET-OUT
4	10.0.0.20	111111:111120	AAA	Eth2	OFPT-PACKET-OUT
5	10.0.2.2	111111:111122	AAA	Eth3	



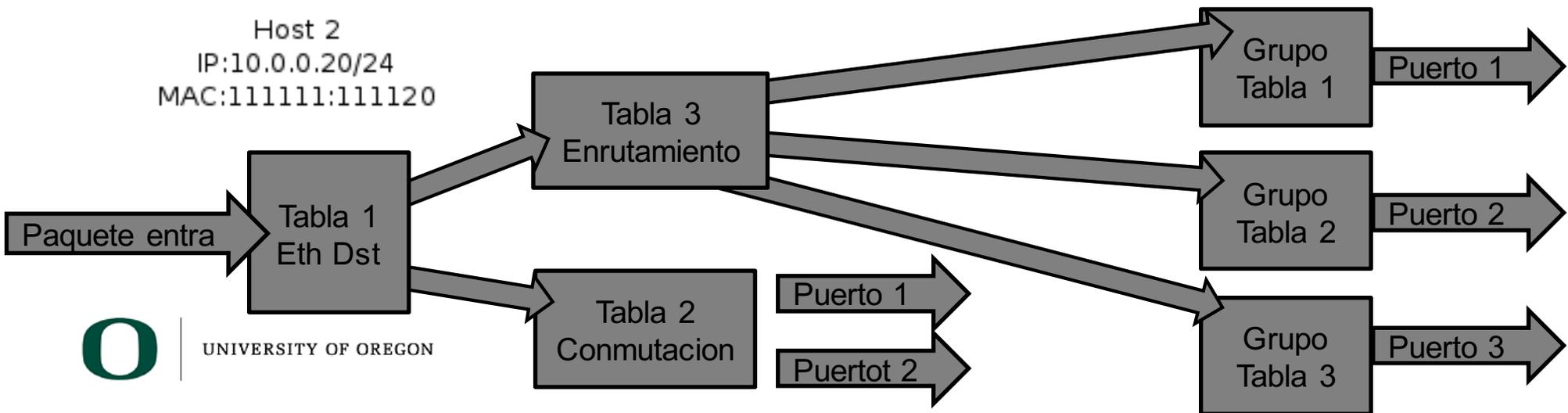
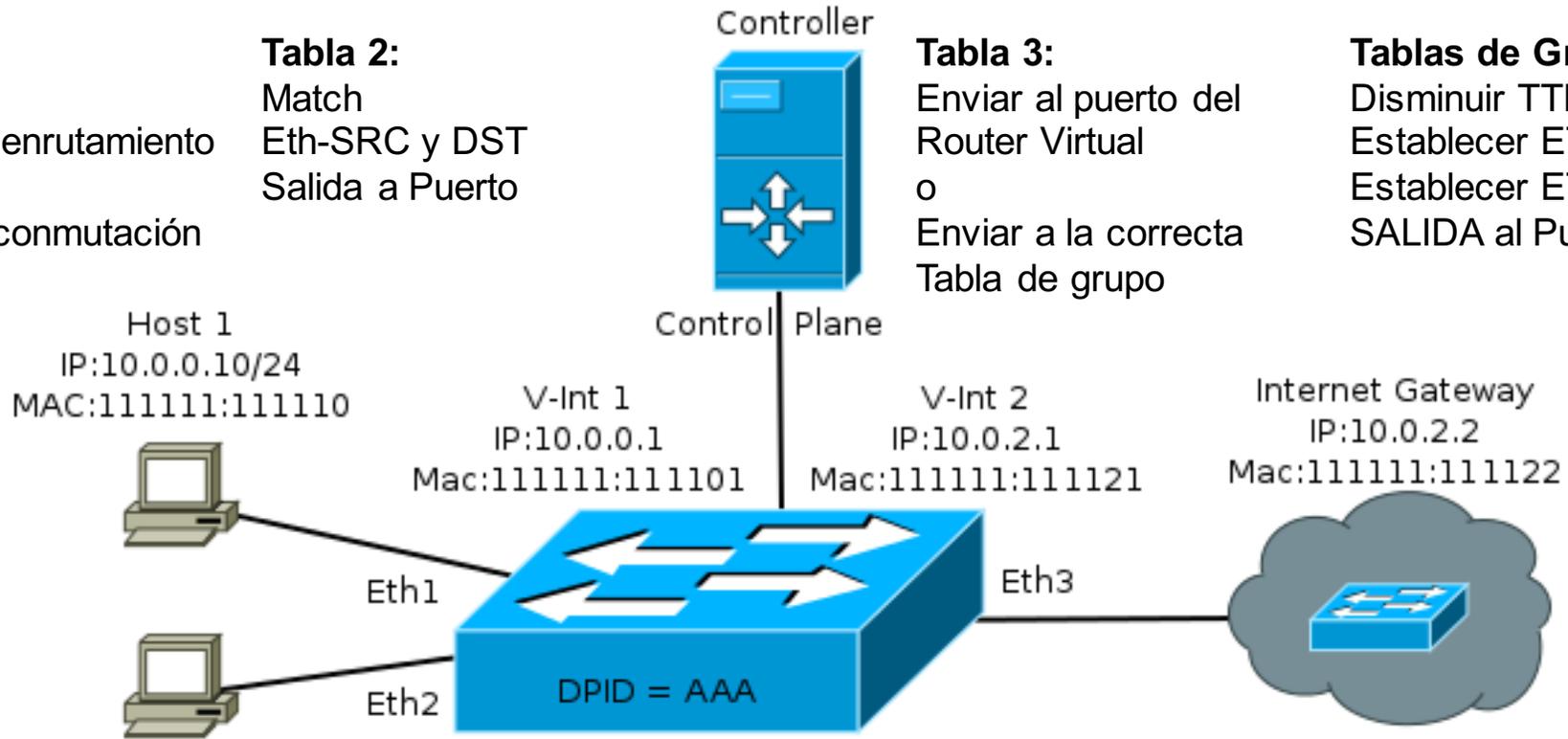
# Tuberia

**Tabla 1:**  
 Enviar a  
 Tabla de enrutamiento  
 o  
 tabla de conmutación

**Tabla 2:**  
 Match  
 Eth-SRC y DST  
 Salida a Puerto

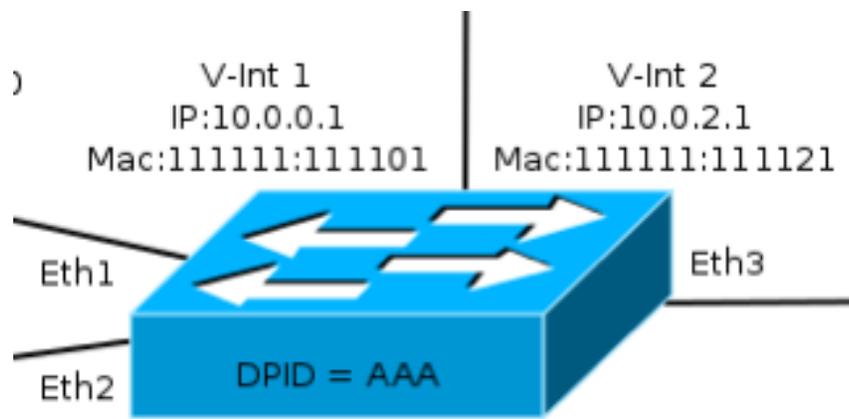
**Tabla 3:**  
 Enviar al puerto del  
 Router Virtual  
 o  
 Enviar a la correcta  
 Tabla de grupo

**Tablas de Grupo:**  
 Disminuir TTL  
 Establecer ETH SRC  
 Establecer ETH DST  
 SALIDA al Puerto



# Tabla 1 (Enrutamiento o Conmutacion)

Table 1		
Ethertype	Ethernet-dst	Action
==ARP	*	OFPAT_OUTPUT to CONTROLLER
*	111111:111101	OFPIT_GOTO_TABLE 2 (Routing)
*	111111:111121	OFPIT_GOTO_TABLE 2 (Routing)
*	*	OFPIT_GOTO_TABLE 3 (Switching)



Si es paquete ARP, enviar al controlador

Si MAC DST es una interfaz del router virtual, Enviar a la tabla de enrutamiento.

De lo contrario, enviar a la tabla de conmutación.

# Tabla de conmutación

Tabla 2 (conmutación)		
Eth-SRC	Eth-DST	Accion
111111:111110	111111:111120	OFPAT_OUTPUT to Eth 2
111111:111120	111111:111110	OFPAT_OUTPUT to Eth 1

Si Eth-SRC es Host 1 y Eth-DST es Host 2, enviar al Puerto Ethernet 2

Si Eth-SRC es Host 2 y Eth-DST es Host 1, enviar al Puerto Ethernet 1



# Tabla de enrutamiento

Tabla 3 (Enrutamiento)		
IP-SRC	IP-DST	Accion
*	10.0.0.1	OFPAT_OUTPUT to LOCAL
*	10.0.2.1	OFPAT_OUTPUT to LOCAL
*	10.0.0.10	OFPAT_GROUP 1
*	10.0.0.20	OFPAT_GROUP 2
*	10.0.2.2	OFPAT_GROUP 3

Si IP-DST es para una interface del Router Virtual, enviar a la de pila local (switch).

Si IP-DST es para el Host 1 envío a la tabla del Grupo 1

Si IP-DST es para el Host 2 envío a la tabla del Grupo 2

Si IP-DST es para la puerta de enlace a Internet envío a la tabla del Grupo 3

# Tabla de Grupos

Tabla de Grupo 1		
No.	Accion	Valor
1	OFPAT_DEC_NW_TTL	-
2	OFPAT_SET_FIELD OXM_OF_ETH_SRC	111111:111101 (V-Int 1)
3	OFPAT_SET_FIELD OXM_OF_ETH_DST	111111:111110 (Host 1)
4	OFPAT_OUTPUT	Eth 1

- 1) Disminuir el TTL
- 2) Establecer ETH SRC para la interface 1 del Router Virtual
- 3) Establecer ETH DST para el Host 1
- 4) Salida al Puerto Eth 1

# Tabla de Grupos (Cont)

Tabla de Grupo 2		
No.	Accion	Valor
1	OFPAT_DEC_NW_TTL	-
2	OFPAT_SET_FIELD OXM_OF_ETH_SRC	111111:111101 (V-Int 1)
3	OFPAT_SET_FIELD OXM_OF_ETH_DST	111111:111120 (Host 2)
4	OFPAT_OUTPUT	Eth 2

Tabla de Grupo 3		
No.	Accion	Valor
1	OFPAT_DEC_NW_TTL	-
2	OFPAT_SET_FIELD OXM_OF_ETH_SRC	111111:111121 (V-Int 2)
3	OFPAT_SET_FIELD OXM_OF_ETH_DST	111111:111122 (ISP GW)
4	OFPAT_OUTPUT	Eth 3

# Ejercicio Multi-tabla



UNIVERSITY OF OREGON

