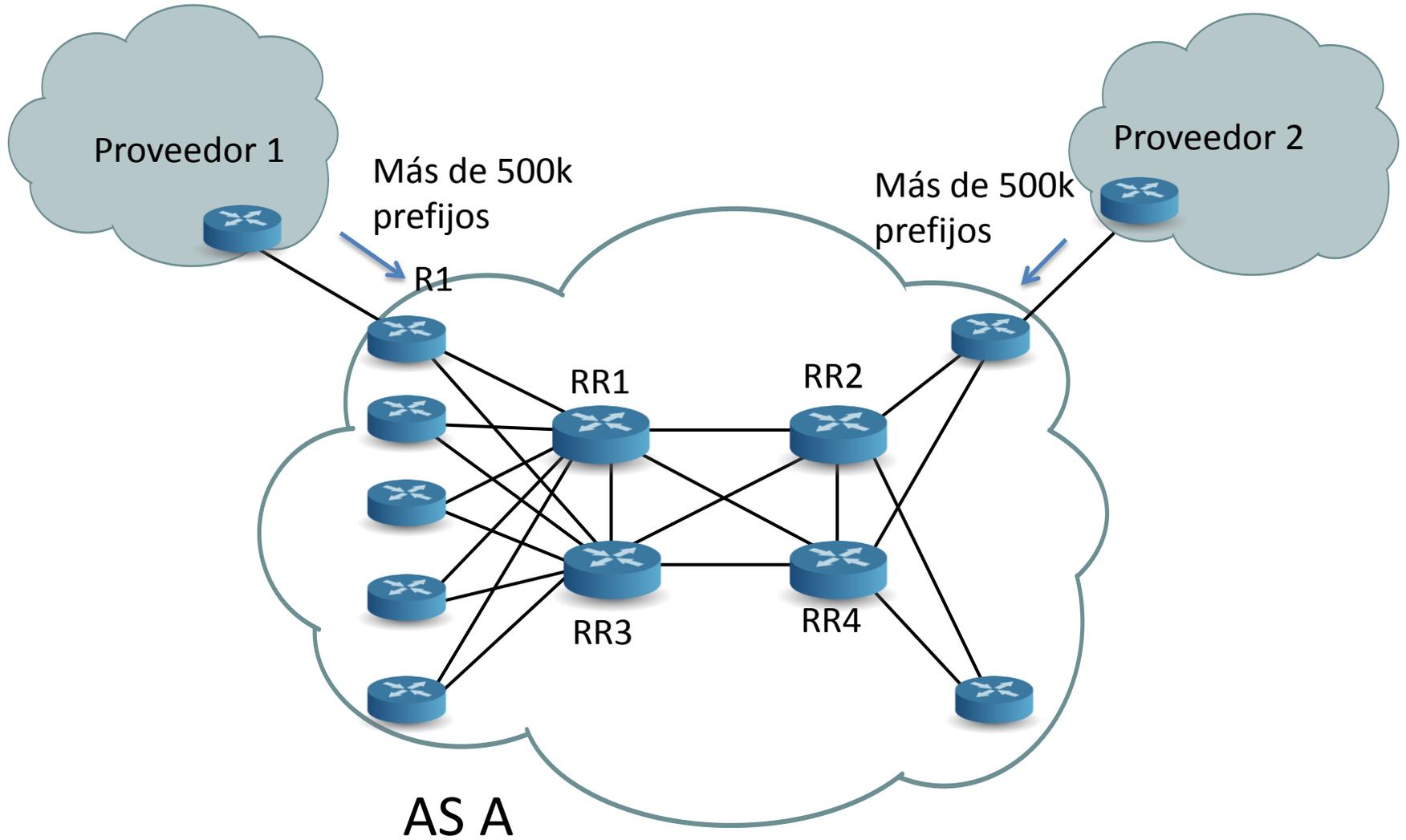


Reduciendo el tráfico perdido durante la reconvergencia de BGP

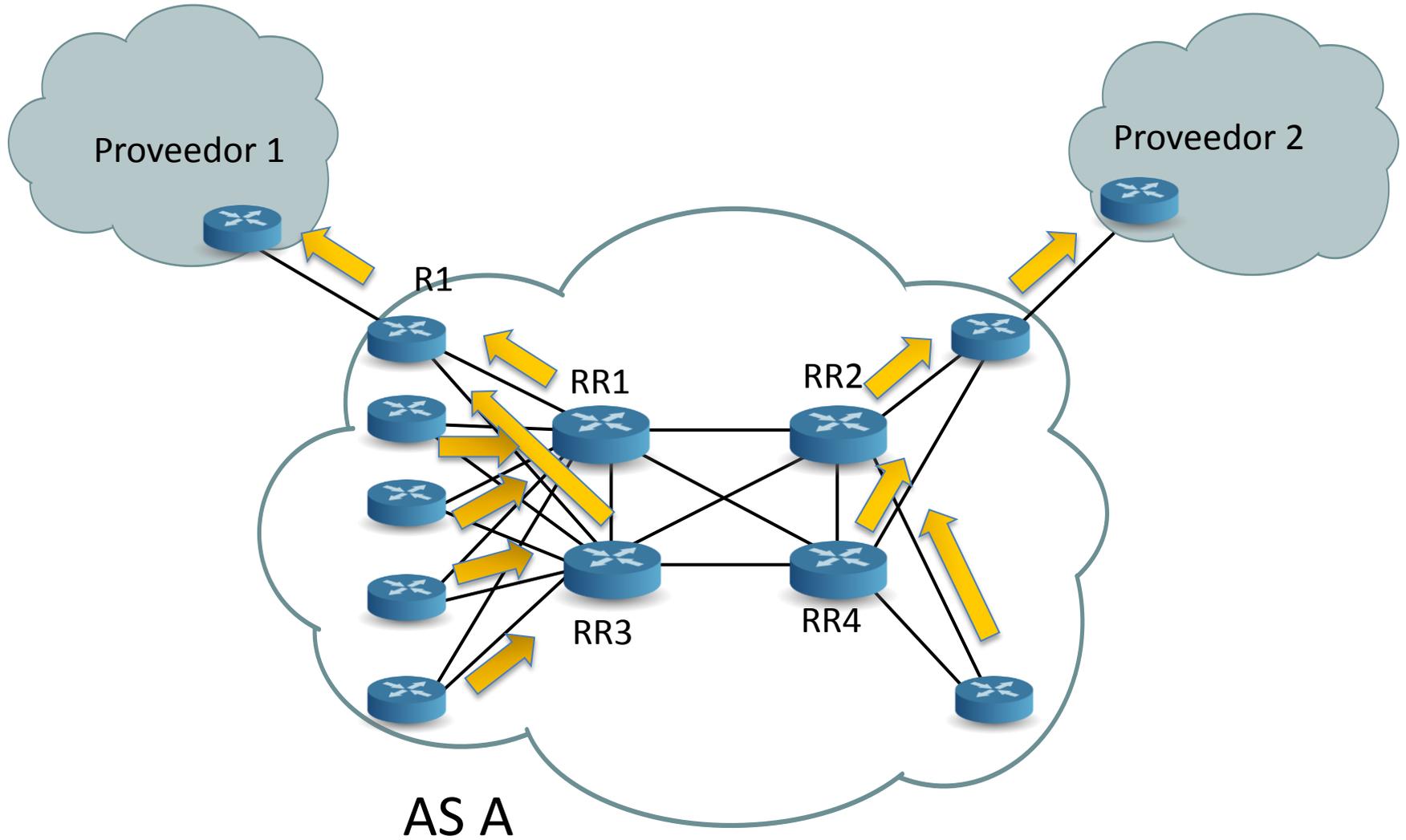
Presentado por Alberto García Martínez (UC3M)
alberto@it.uc3m.es

Trabajo realizado en colaboración con Andra
Lutu (Simula), Marcelo Bagnulo (UC3M), Cristel
Pelsser (IIJ)

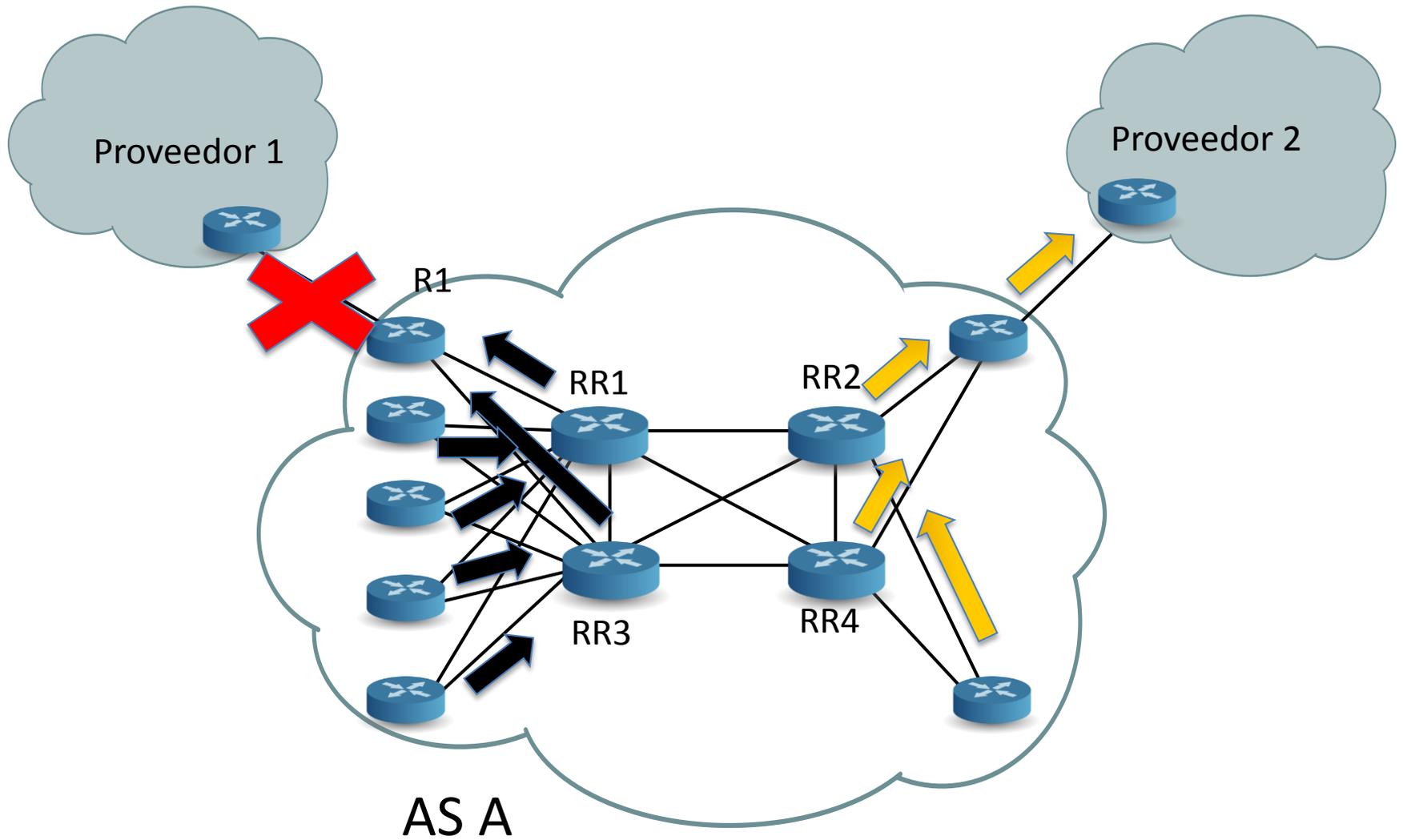
Punto de partida



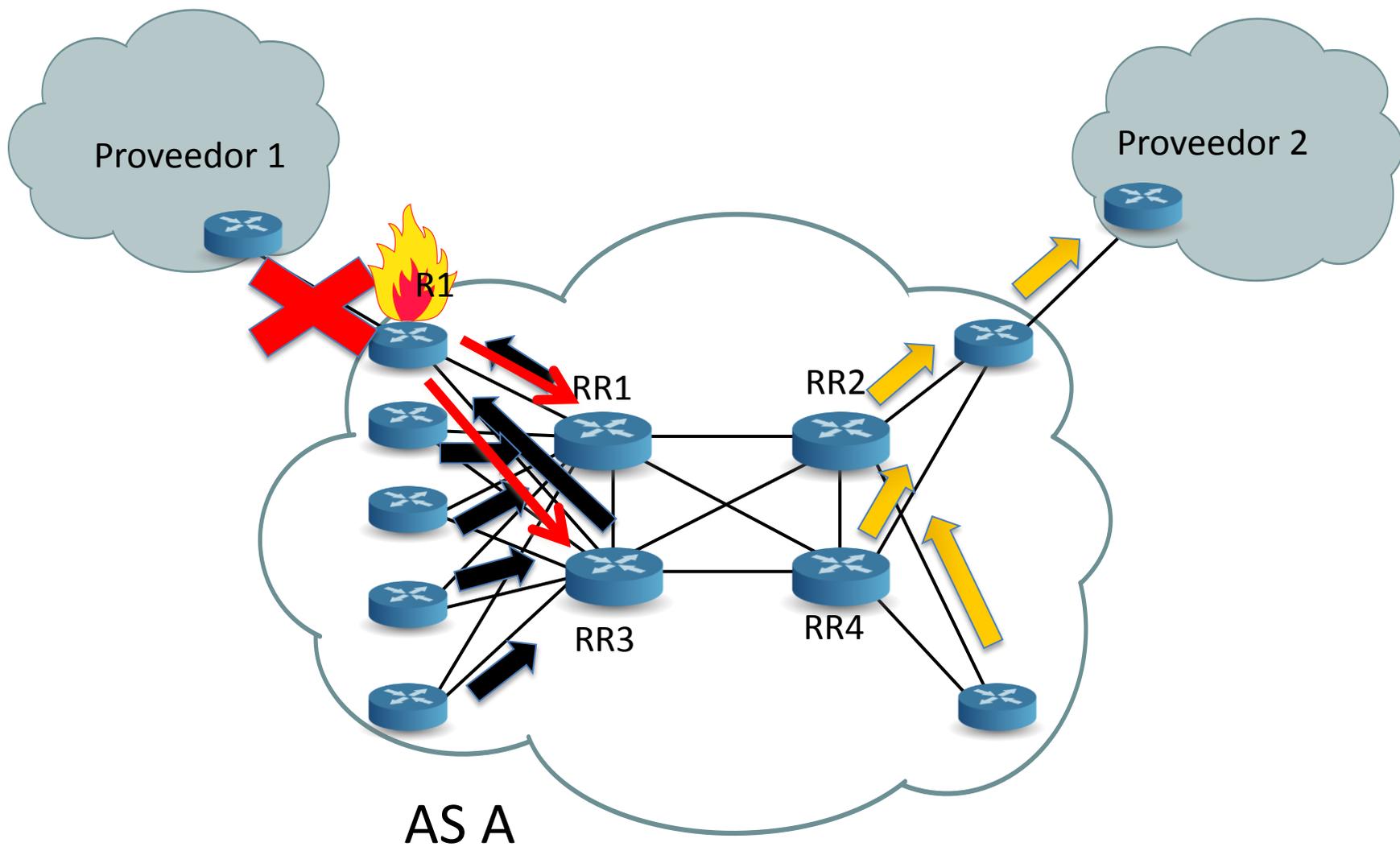
Punto de partida



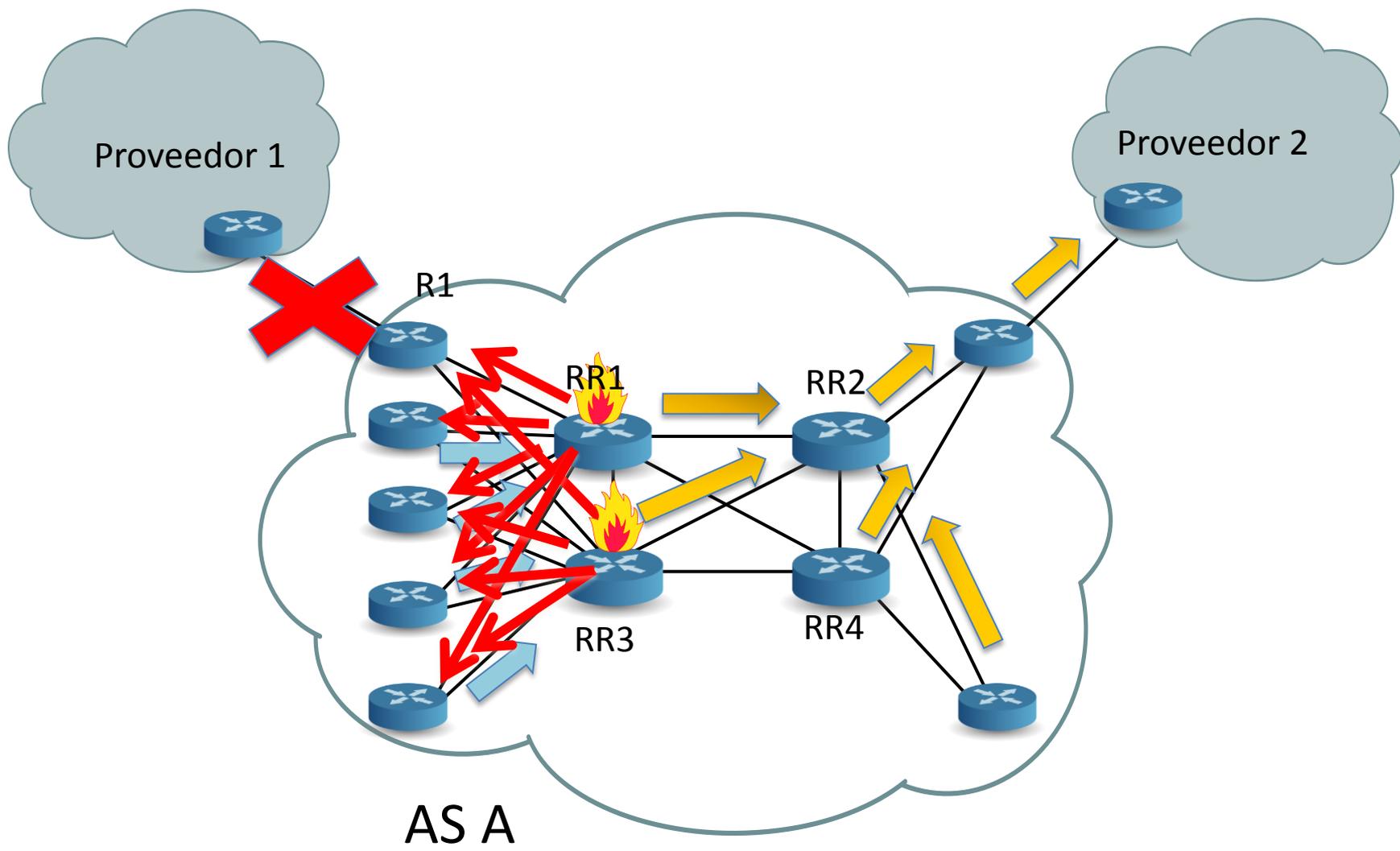
El Problema



El Problema



El Problema



El Problema

- R1 tarda varios segundos en converger
 - Ejecutando BGP con Quagga, del orden de 30 s en una topología como la figura
 - Puede tardar más si hay más sesiones BGP
- Mientras, se pierden paquetes (multiplicar tiempo por *rate*)

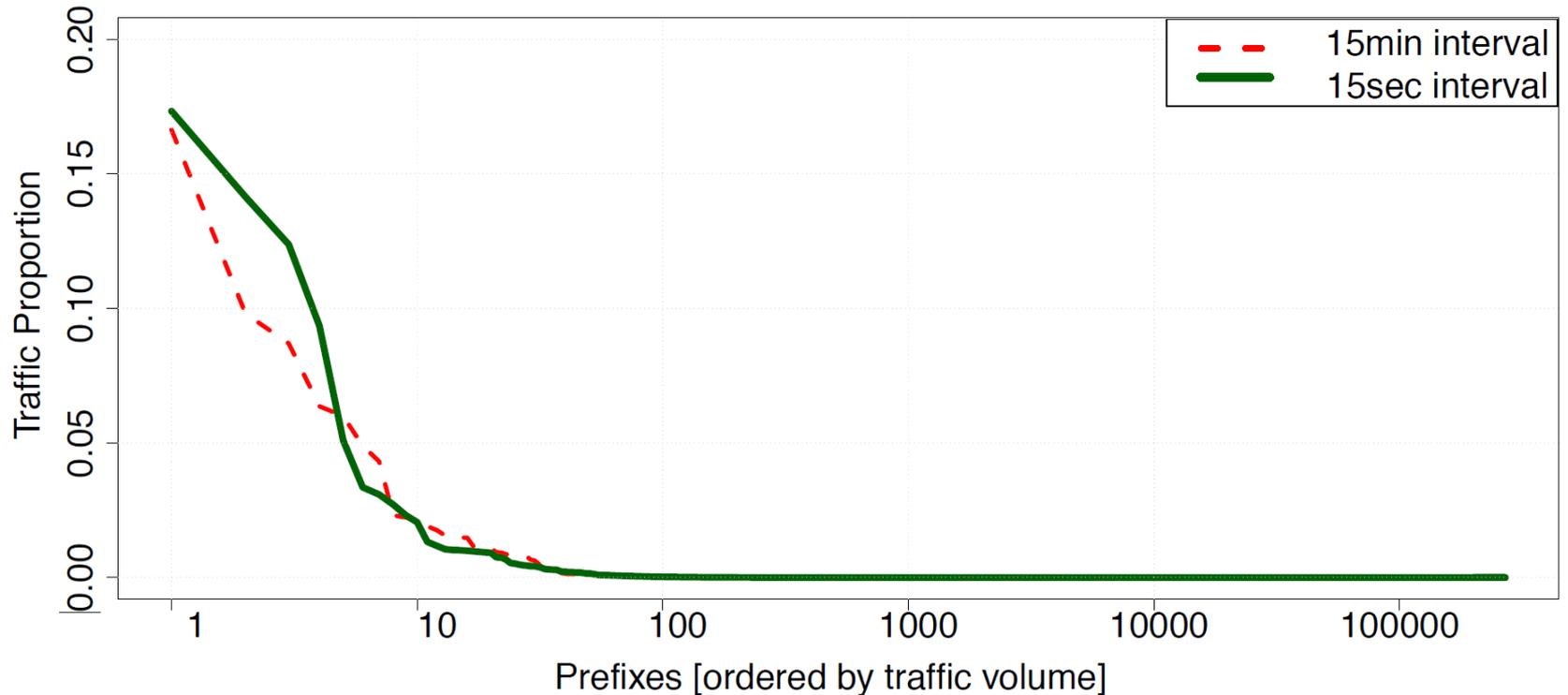
Observación 1: Unos pocos destinos BGP reciben la mayoría del tráfico

Comcast's [Brian Field](#) shared one very interesting observation: in a six day period in their network they observed that some 415K entries had no traffic at all! Some 90% of the data traffic handled by the routers was directed to 3,156 distinct routing prefixes, and 99% of the traffic was sent to 25,893 prefixes.

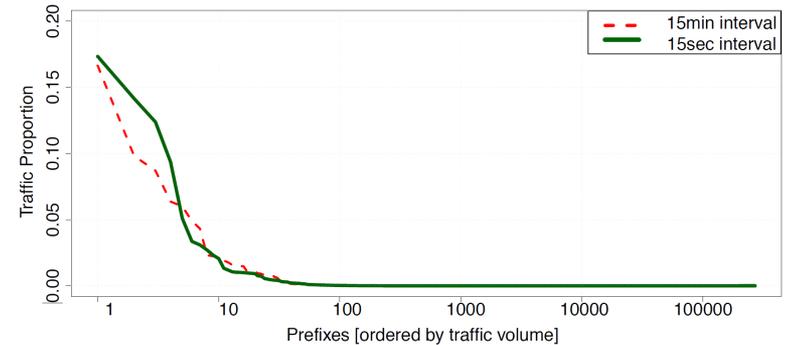
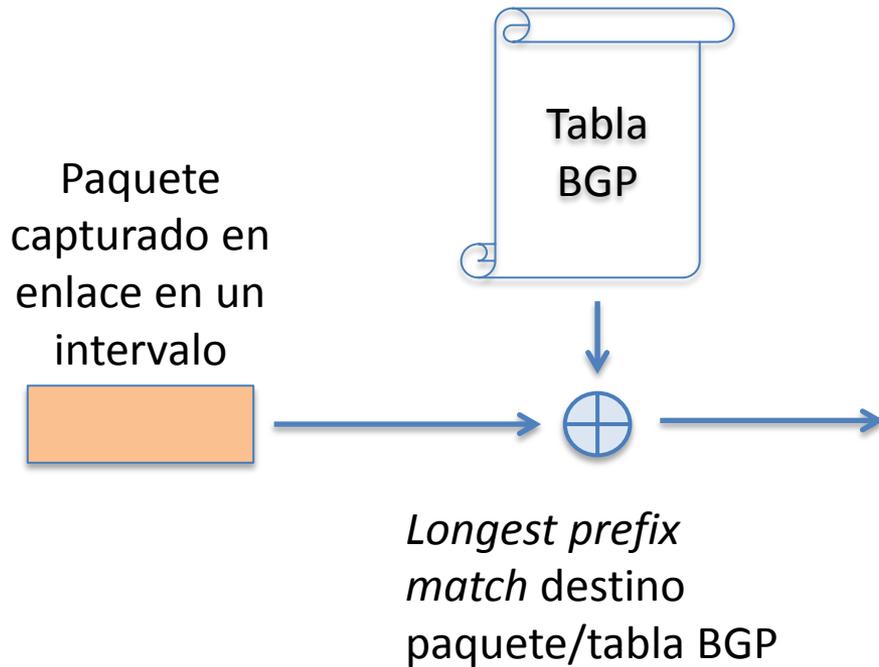
Geoff Huston, jun 2015

<https://blog.apnic.net/2015/06/05/thoughts-from-nanog-64/>

Observación 1: Unos pocos destinos BGP reciben la mayoría del tráfico

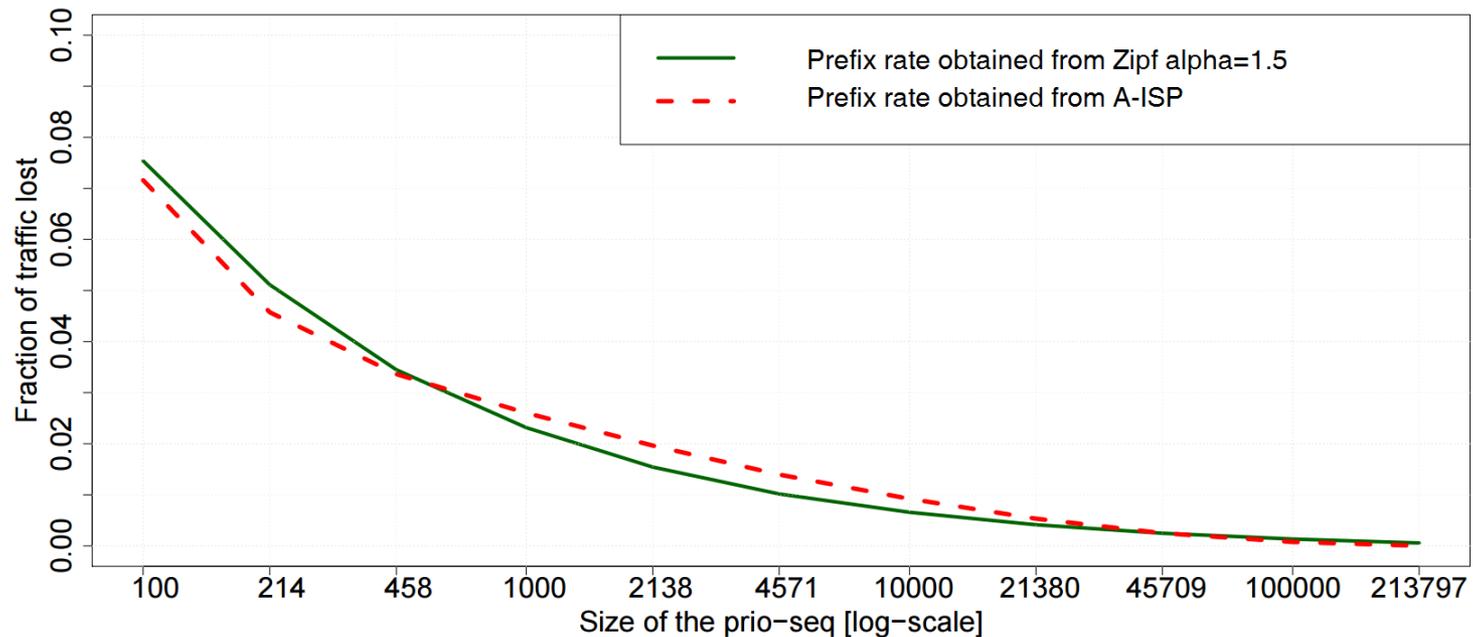


Observación 1: Unos pocos destinos BGP reciben la mayoría del tráfico



Observación 1: Unos pocos destinos BGP reciben la mayoría del tráfico

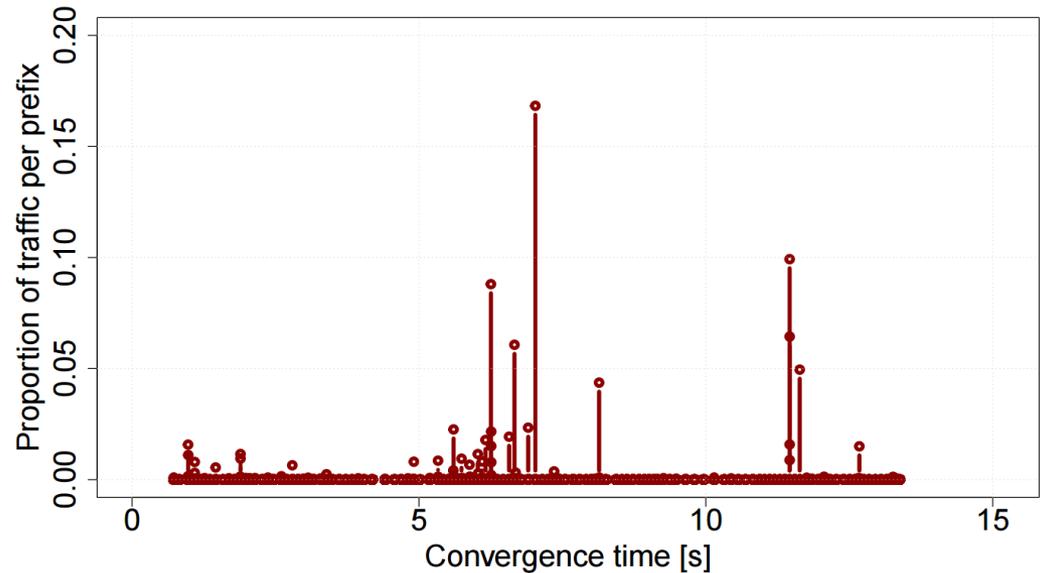
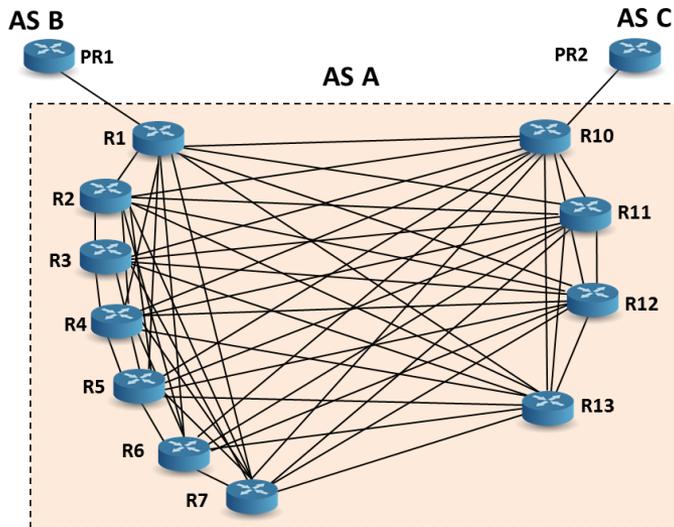
- Distribución de tráfico de tipo *power-law*, posiblemente Zipf



Observación 2: hay eventos singulares que afectan a un gran número de rutas

- Ejemplo de eventos singulares que afectan al recálculo de un gran número de rutas
 - Caída de un enlace con un proveedor
 - Hay que borrar todas las rutas que se recibieron del proveedor, calcular nuevas, mandar rutas mejores a otros
 - Nueva conexión BGP establecida con un cliente
 - Hay que mandar todas las rutas propias
 - Cambio de políticas en un router
 - Cambio en la información del IGP (puede afectar a la alcanzabilidad de las rutas)

Observación 3: Los prefijos afectados por eventos singular tardan un tiempo diferente en reconverger



Idea

Cuando ocurra un evento singular,

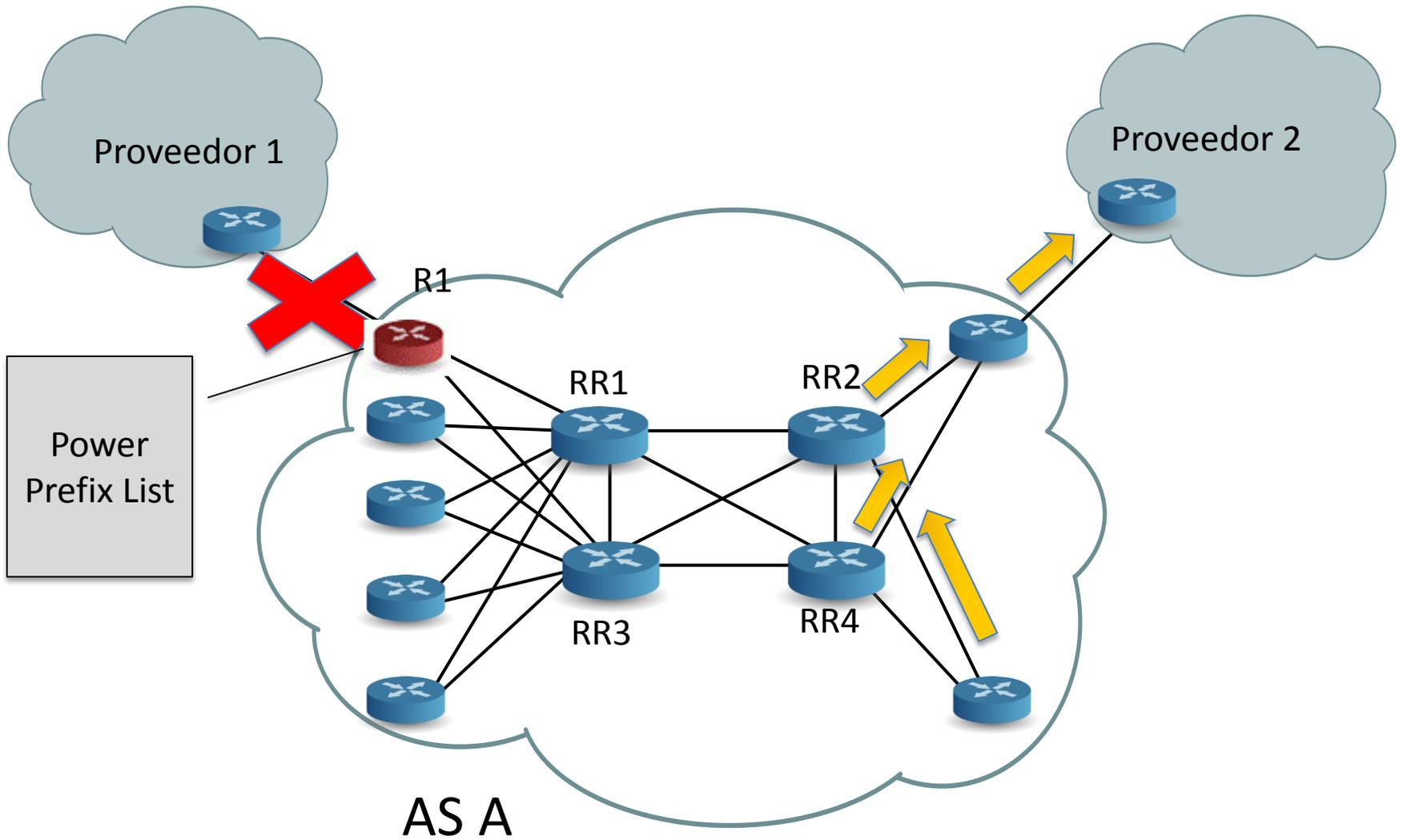
ordenar el procesado de rutas BGP

según la contribución de cada prefijo al tráfico

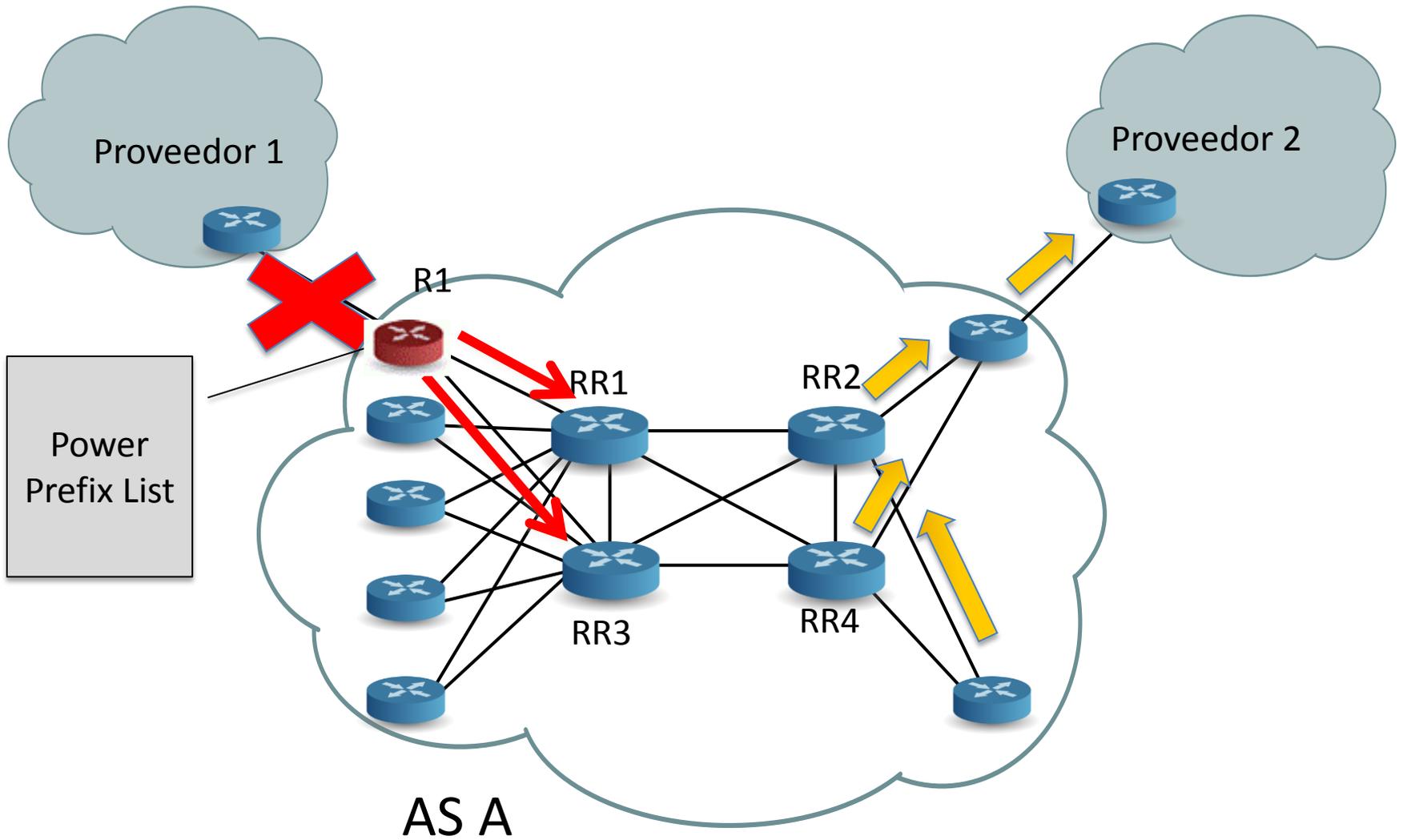
para reducir la cantidad de tráfico perdida en uno de estos eventos



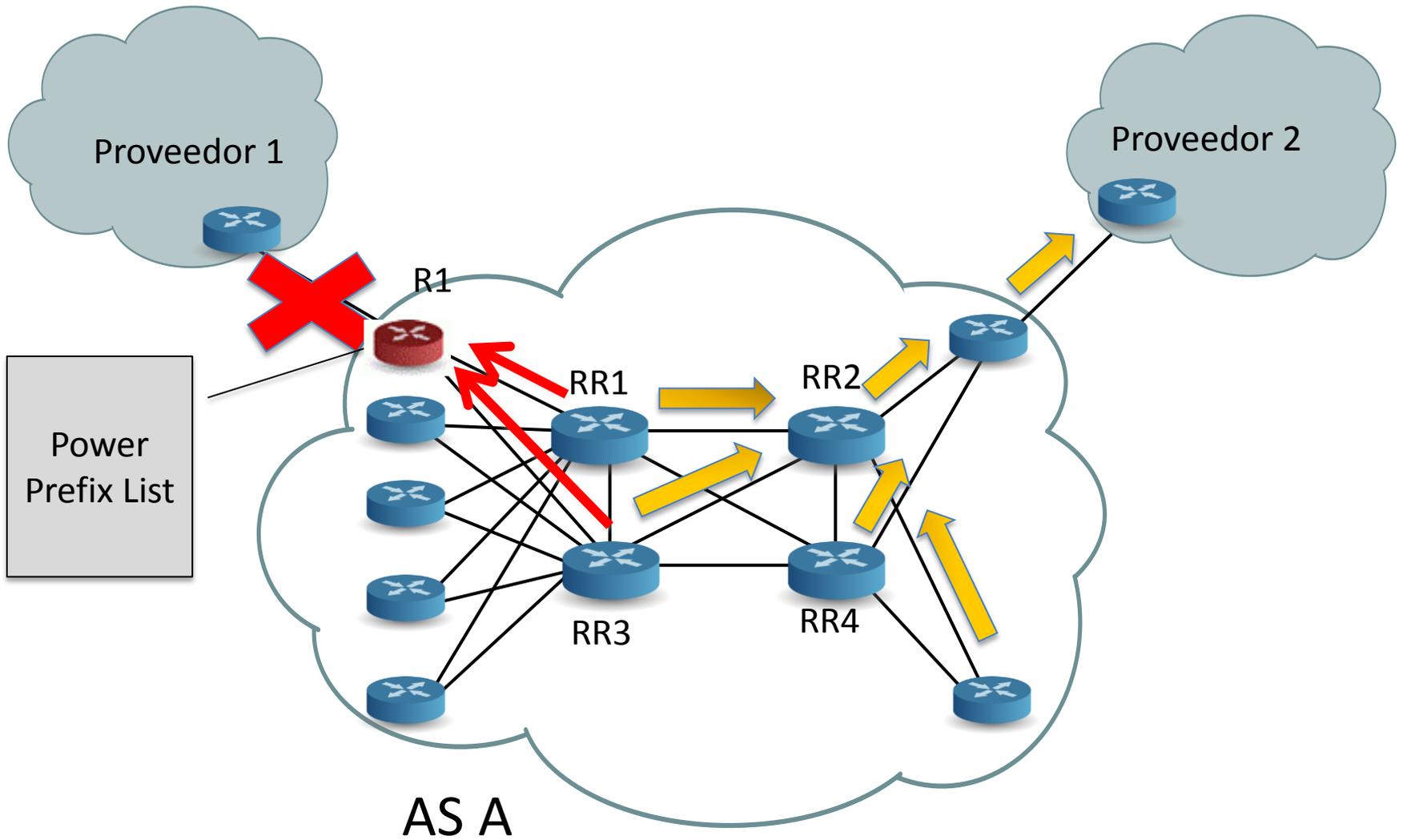
Solución



Solución



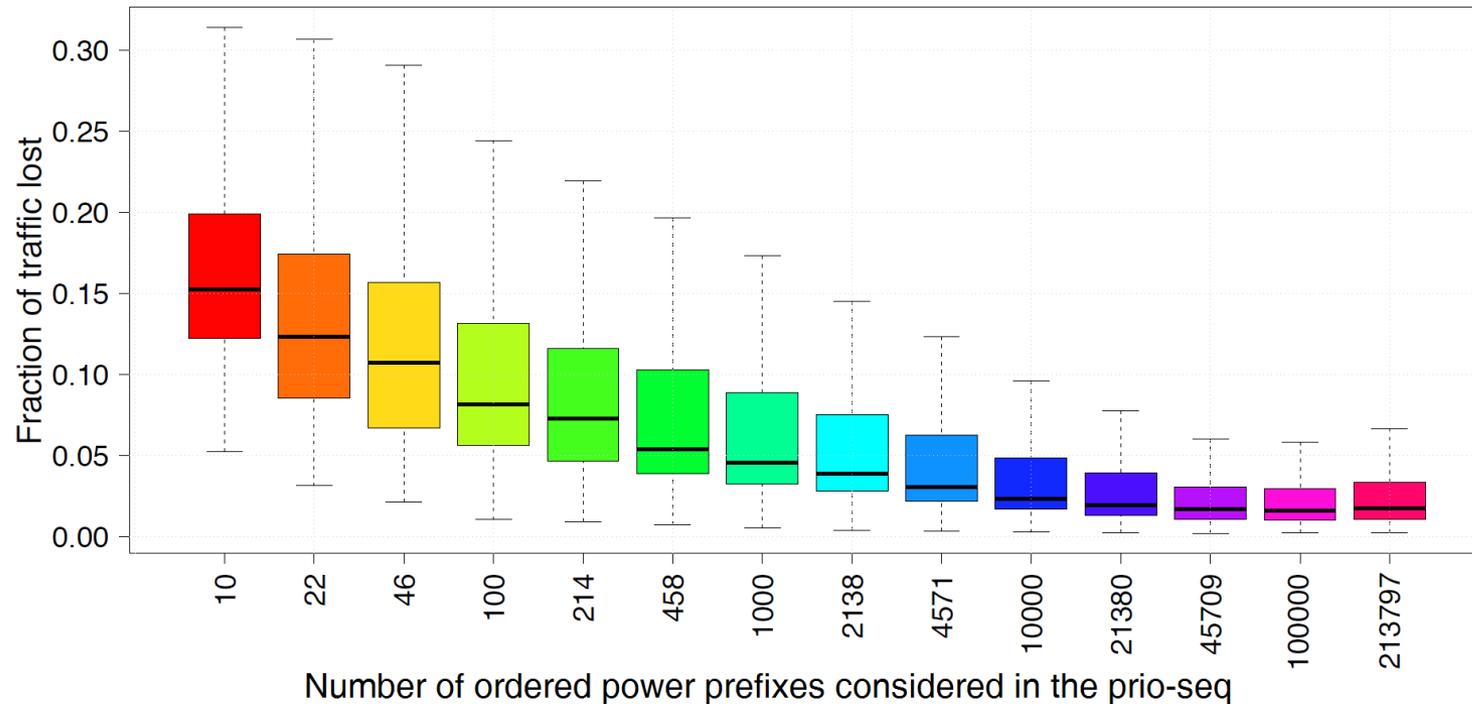
Solución



Algunos flecos

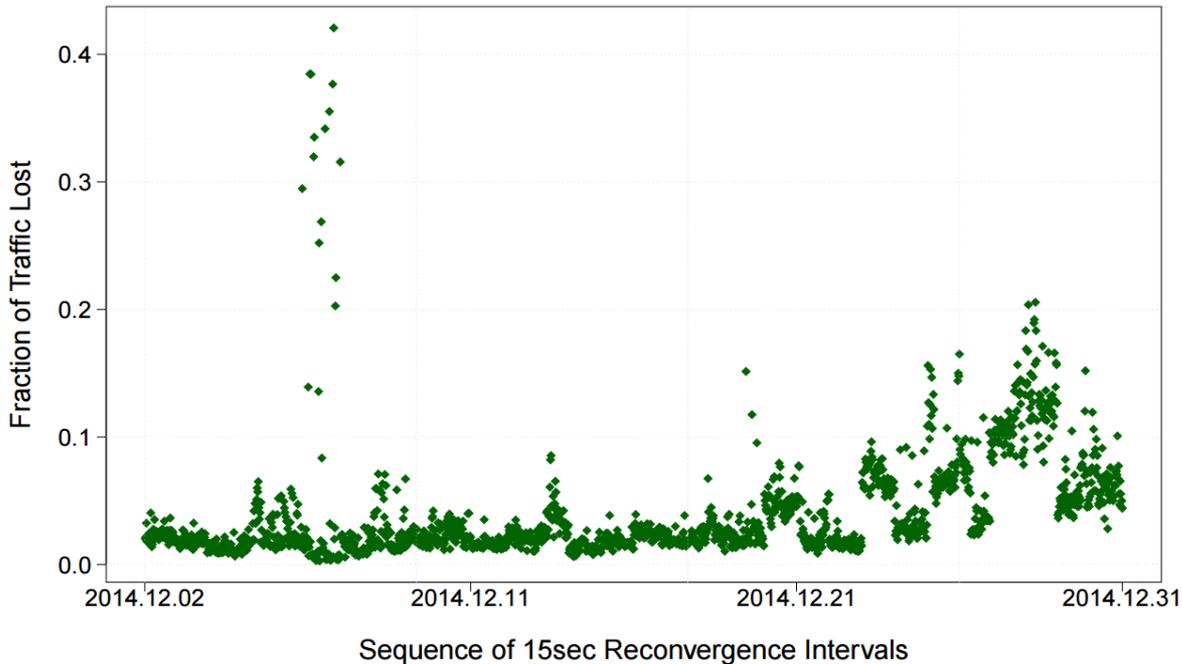
- Objetivo:
 - Que la lista prediga lo mejor posible el tráfico durante periodos de reconvergencia (~decenas de segundos)
 - No muy costoso de gestionar
- Para eso hay que definir
 - ¿Cuántos prefijos tiene que tener la lista priorizada en los routers?
 - ¿Cómo generar la lista?
 - ¿Durante cuánto tiempo hay que medir?
 - ¿En qué instante hay que medir?
 - ¿Durante cuánto tiempo es válida la lista?
- Analizamos datos de un proveedor para responder a estas preguntas

¿Cuántos prefijos?



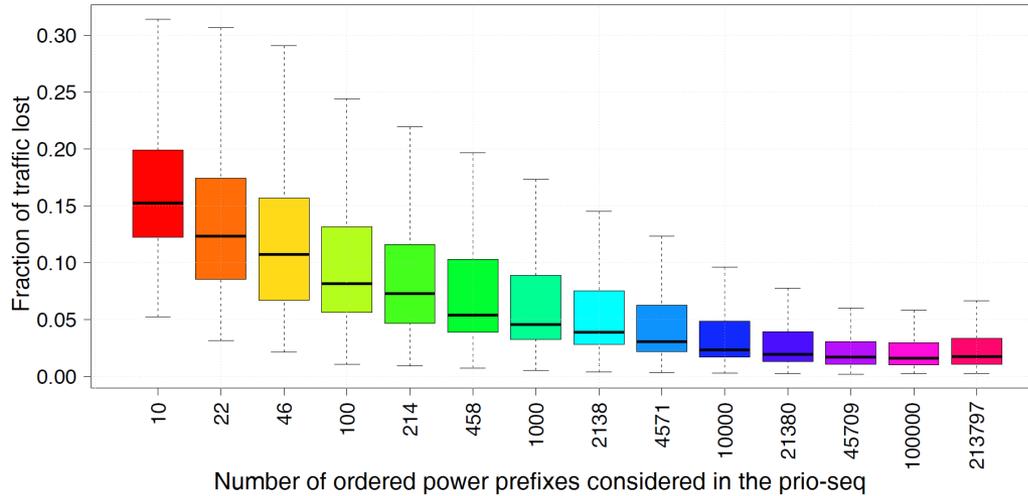
- Predictor = tráfico medido 1 dic 2014, 14:00 a 14:15
- Analizamos cómo predice el tráfico perdido en un conjunto de muchos intervalos de 15 segundos a lo largo de dic de 2014
- ~45000 prefijos es ok

¿Validez del predictor?

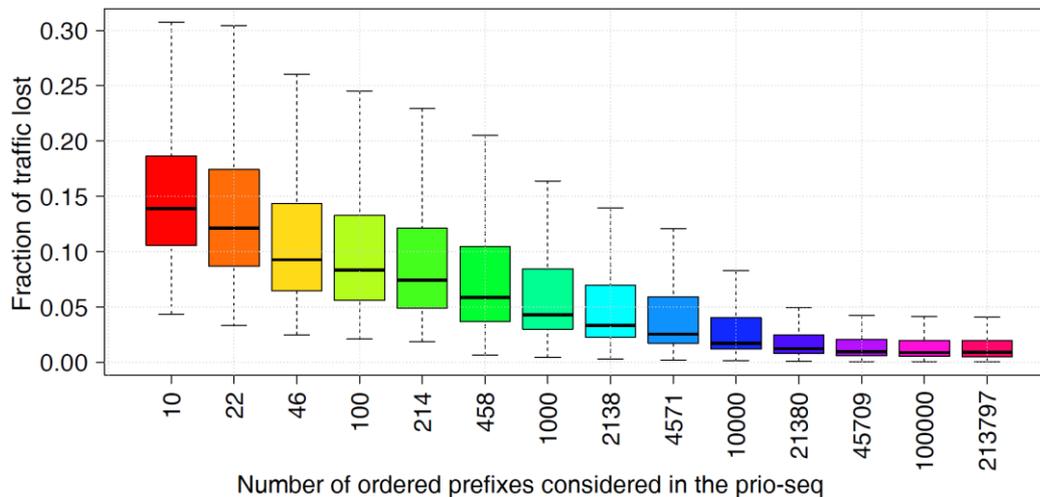


- A medida que el tiempo pasa, la predicción es peor
- Recalcular el predictor cada 15 días parece razonable

¿Cuánto/cuándo medir?



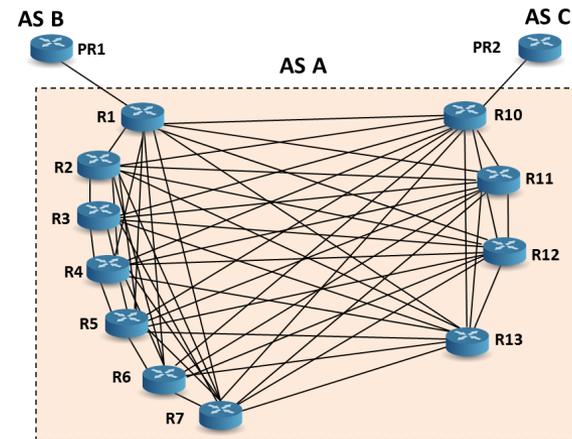
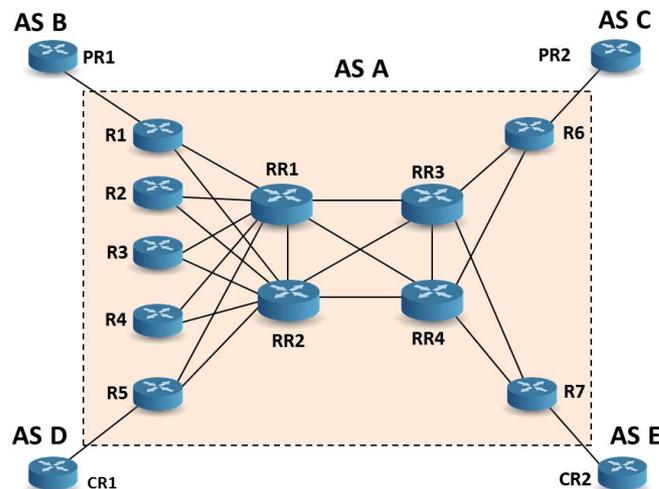
Predictor generado con
15 minutos de tráfico



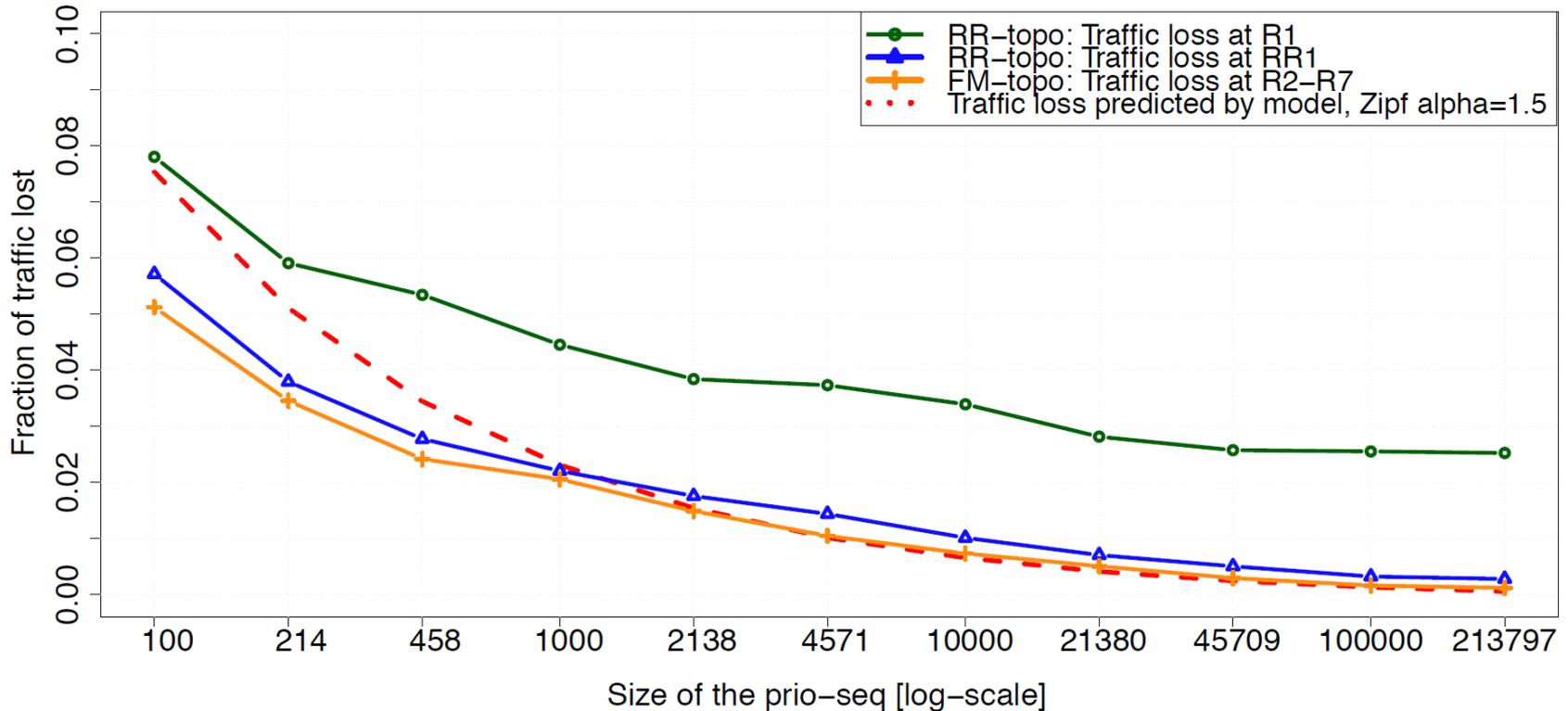
Predictor generado con
1 día de tráfico

Experimento

- Modificamos Quagga 0.99.23 para ordenar el tratamiento de los prefijos según una lista
- PR1 y PR2 inyecta los mismos prefijos
- Utilizamos distribución tráfico igual a la utilizada en el predictor (= el predictor es perfecto en este experimento)
- Tiramos enlace entre PR1 y R1, y medimos tráfico perdido



Resultados experimentales



- Tráfico perdido desde recepción del primer paquete en el router considerado
- Tráfico total perdido es superior, aunque siempre menor que el 23% del total

Conclusiones

- Ordenar procesamiento de prefijos BGP permite reducir significativamente la cantidad de paquetes perdidos
- Fácil despliegue
 - Solo requiere
 - Actualización del software BGP del router
 - Cada cierto tiempo, proceso de medida y configuración del router
 - Podría ser automático en el router, o cargado externamente
 - No requiere
 - Modificación del protocolo BGP
 - Se puede ir poniendo en routers de forma independiente