

Anycast

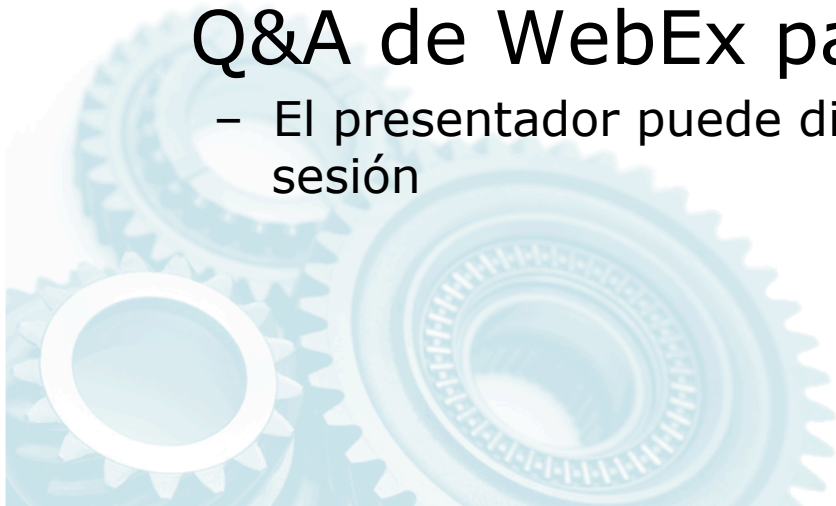
Introducción y Experiencia Operativa

*Presentado por
Luis E. Muñoz, Ingeniero de Aplicaciones*



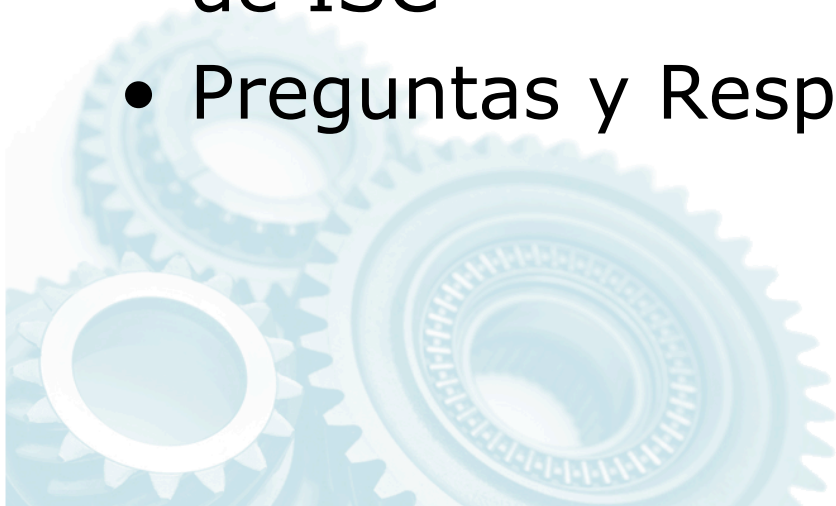
Logística

- La presentación durará 1 hora
- La grabación estará disponible en unos días
 - <http://www.isc.org/webinars>
- Los participantes no intervienen oralmente
- Las preguntas se ingresan vía el panel Q&A de WebEx para el presentador
 - El presentador puede diferir preguntas hasta el final de la sesión



Agenda

- Definición de Anycast
- Explorando los Casos de Uso
- Explorando el Impacto en los Protocolos
- Explorando Anycast y DNS
- Compartiendo la Experiencia Operacional de ISC
- Preguntas y Respuestas



Definición

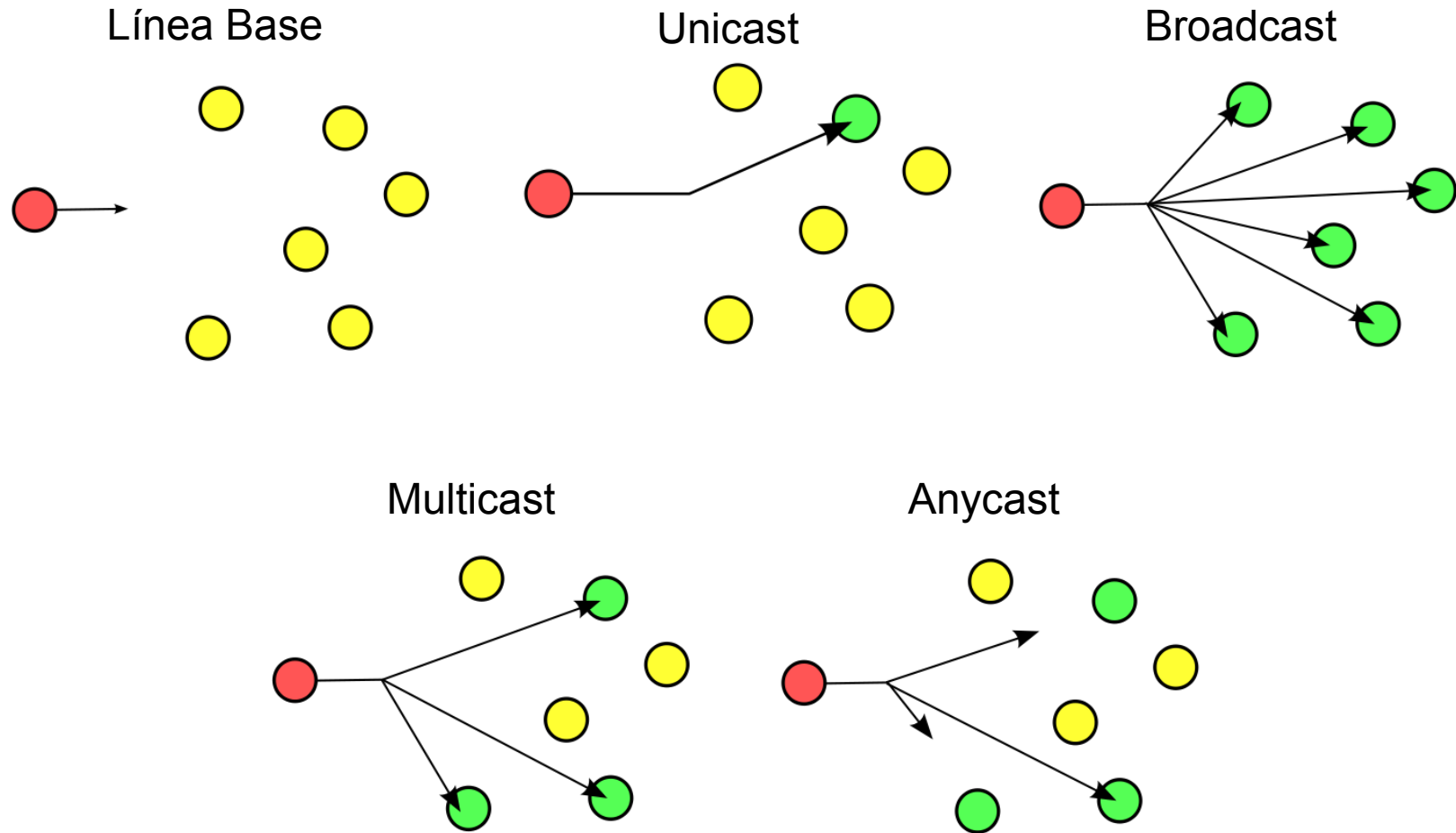
ANYCAST



¿Qué es Anycast?

- Anycast describe un método para usar la misma dirección IP en múltiples servidores
- Anycast es fundamentalmente un *esquema de enrutamiento*
- Anycast tiene más que ver con la configuración de los enrutadores que con la de los servidores
 - Los administradores de los servidores deben estar al tanto de lo que ocurre para poder operar el servicio correctamente

Routing Schemes Compared



Propiedades de Anycast

- Cada paquete enviado a una dirección Anycast puede llegar a un servidor distinto
- Los paquetes van a la dirección con la mejor **métrica**
 - Usualmente el más cercano, pero no siempre. Las métricas pueden basarse en otros factores como ancho de banda, costo, carga o confiabilidad
- Los servidores deben tener una dirección Unicast y una Anycast
 - Las funciones de gestión no se pueden hacer sobre la dirección Anycast porque sólo llegarían a un servidor

Explorando

CASOS DE USO



Use Cases

- Anycast Local

- Distribuye la carga entre múltiples servidores de la misma subred
- Elimina la necesidad de un balanceador al hacer que el enrutador haga la distribución de tráfico

¡Una sola ruta!
Simplifica la operación

A → 1
A → 2
A → 3
A → 4

Enrutamiento
ECMP basado
en flujos

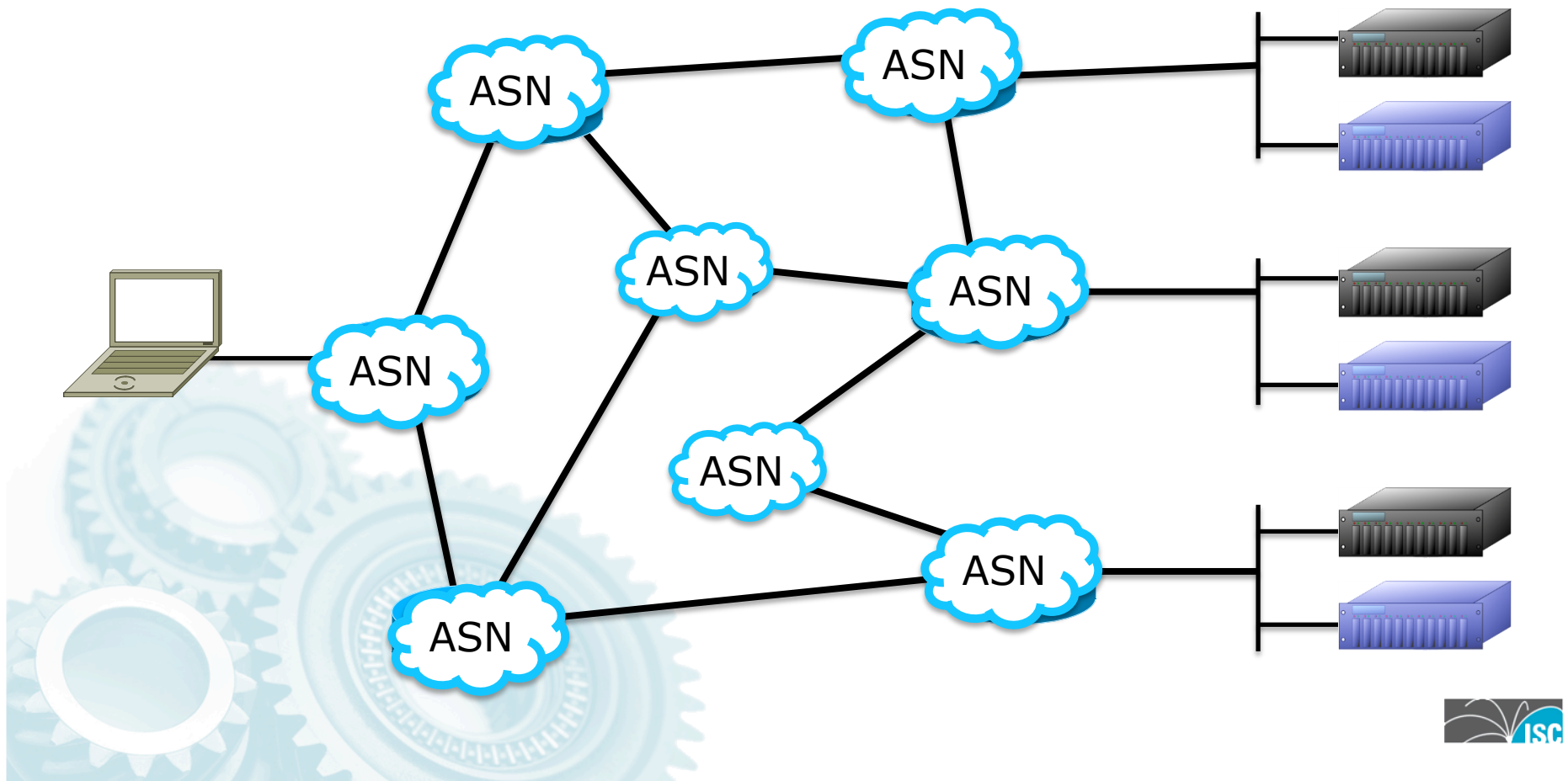


Las rutas se pueden originar con cualquier protocolo

- Estática/RIP/OSPF/ISIS/EIGRP/BGP
- Enrutamiento dinámico para manejar la mayoría de las fallas
- El enrutador puede monitorear los servicios

Anycast Global

- Distribuye la carga entre múltiples localidades
- Ofrece redundancia



Casos de Uso

- Lo más popular en Anycast:

1. DNS, servidores recursivos

- Configurados vía dirección IP en los clientes
- La latencia importa
- Distribuye la carga entre múltiples dispositivos

2. DNS, servidores de autoridad

- Número limitado de direcciones en un sólo paquete de respuesta
- La latencia al servidor es importante
- La redundancia es vital
- Distribución de carga entre dispositivos

3. NTP

- Usualmente sólo en ISPs que tienen grandes cantidades de CPEs que necesitan NTP configurado por dirección IP o suficientes clientes para justificar balanceo de carga

4. Redirectores HTTP

- Servidores HTTP que envían usuarios a otras instancias locales

Explorando

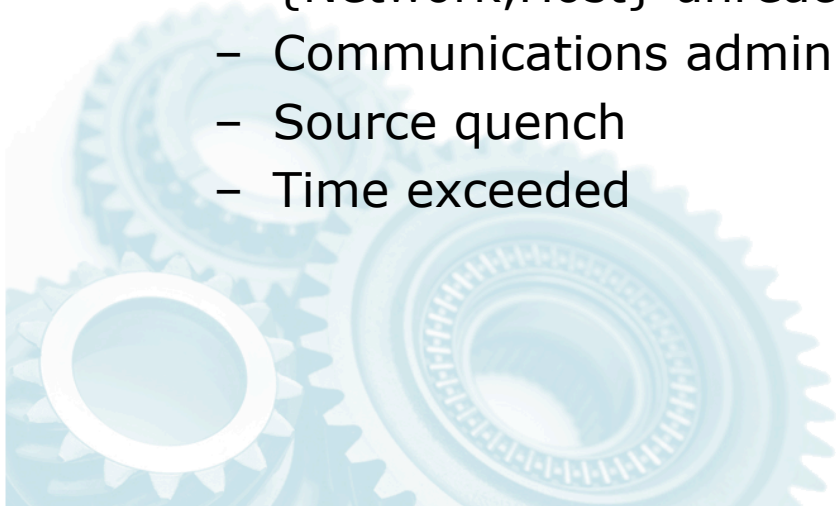
IMPACTO EN LOS PROTOCOLOS

Impacto en Protocolos: ICMP

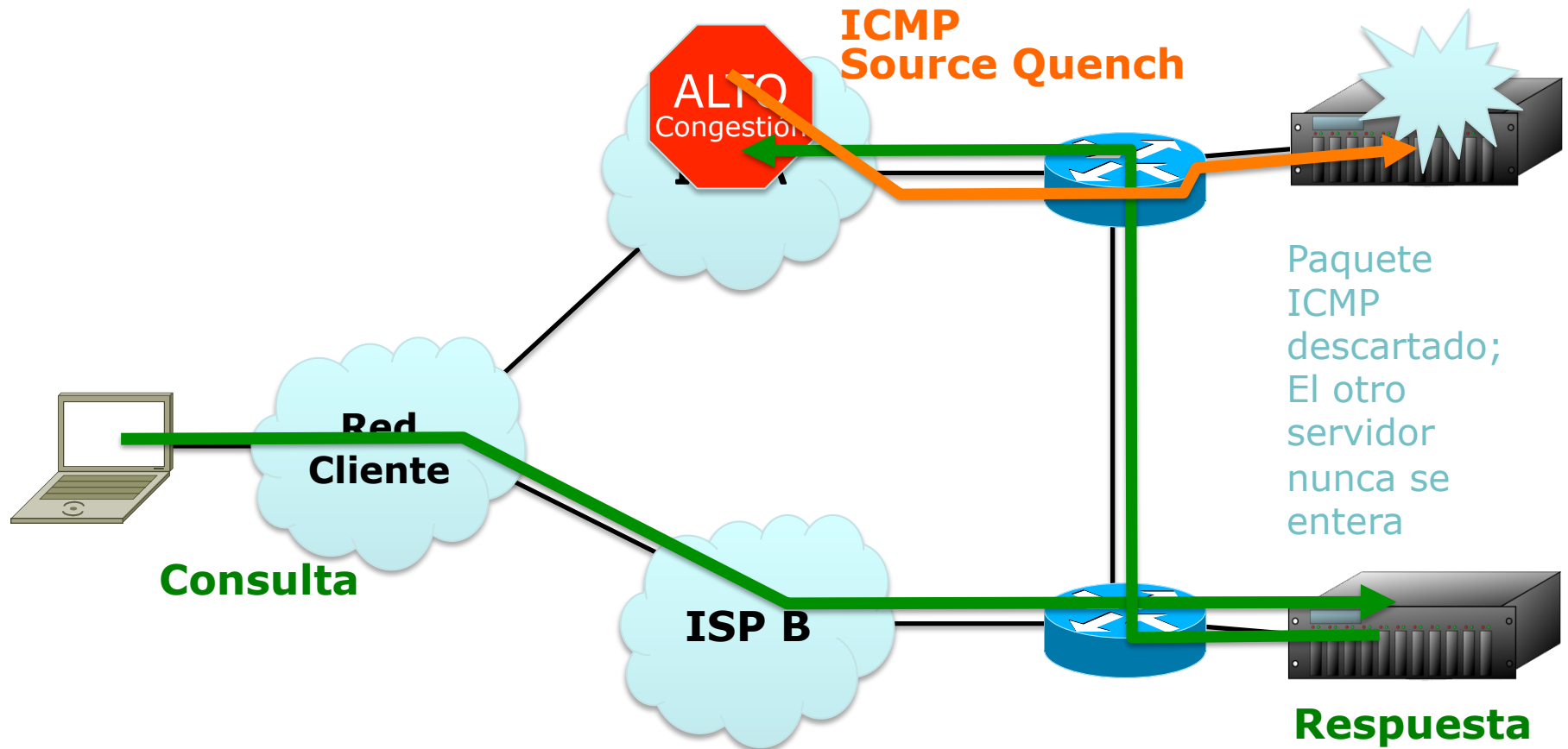
- Las opciones globales y sin estado – *stateless* – funcionan bien
 - Ping solicitud y respuesta
 - ICMP Traceroute
 - La inestabilidad en la red puede causar efectos extraños en traceroute
- Evite las opciones LAN
 - Solicitud y anuncio de enrutadores
 - Solicitud y respuesta de la máscara de red
 - Redirección
 - Una dirección unicast en el servidor puede mitigar estos problemas
 - Es fácil evitar todas esas opciones de ICMP

Impacto en Protocolos: ICMP

- Los mensajes que indican fallas son un problema
 - Destination {network,host,protocol,port} {unreachable,unknown}
 - Fragmentation required
 - Source route failed
 - Source host isolated
 - Network administratively prohibited
 - {Network,Host} unreachable for TOS
 - Communications administratively prohibited
 - Source quench
 - Time exceeded



Impacto en Protocolos: ICMP

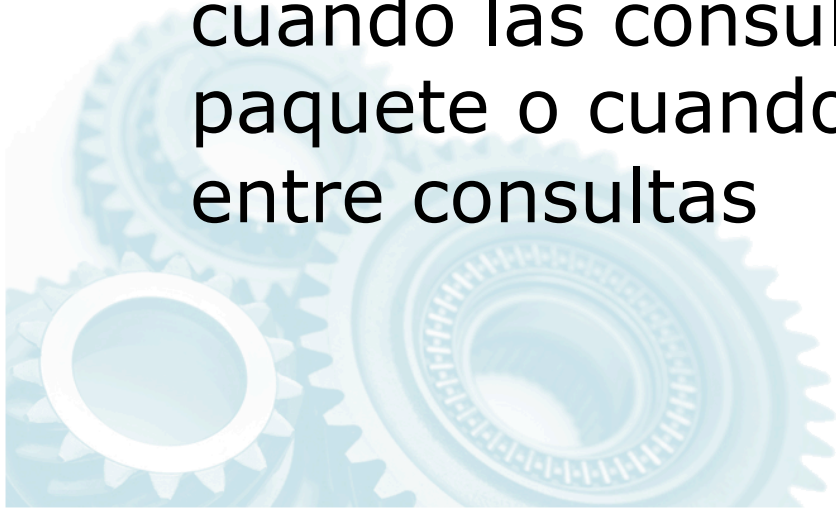


Impacto en Protocolos: ICMP

- Importancia para la operación
- Perder un mensaje “packet too big” previene PMTU
 - Las respuestas de un servidor Anycast **nunca** deben tener el bit DF
 - Se debe asumir fragmentación en el camino o limitar al MTU mínimo
 - **IPv6 no permite fragmentación en el camino.** Todos los paquetes deben respetar el MTU mínimo de 1280 octetos
- La pérdida de mensajes interfiere el cierre de conexiones
 - *Timeouts* para los usuarios finales – puede ser una pausa larga
 - Más consumo de recursos en los servidores que esperan el cierre de las conexiones

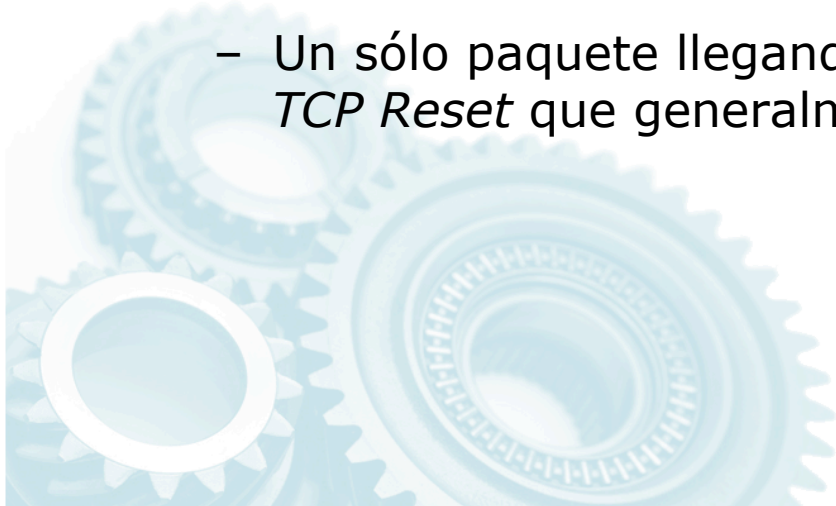
Impacto en Protocolos: UDP

- Sin estado, buen candidato para Anycast
- Funciona bien cuando la consulta va en un paquete y la respuesta en 1-n paquetes, sin estado entre consultas
 - ¡Se parece mucho a la mayoría de las consultas DNS!
- Mismo comportamiento que en TCP cuando las consultas son de más de un paquete o cuando se comparte estado entre consultas

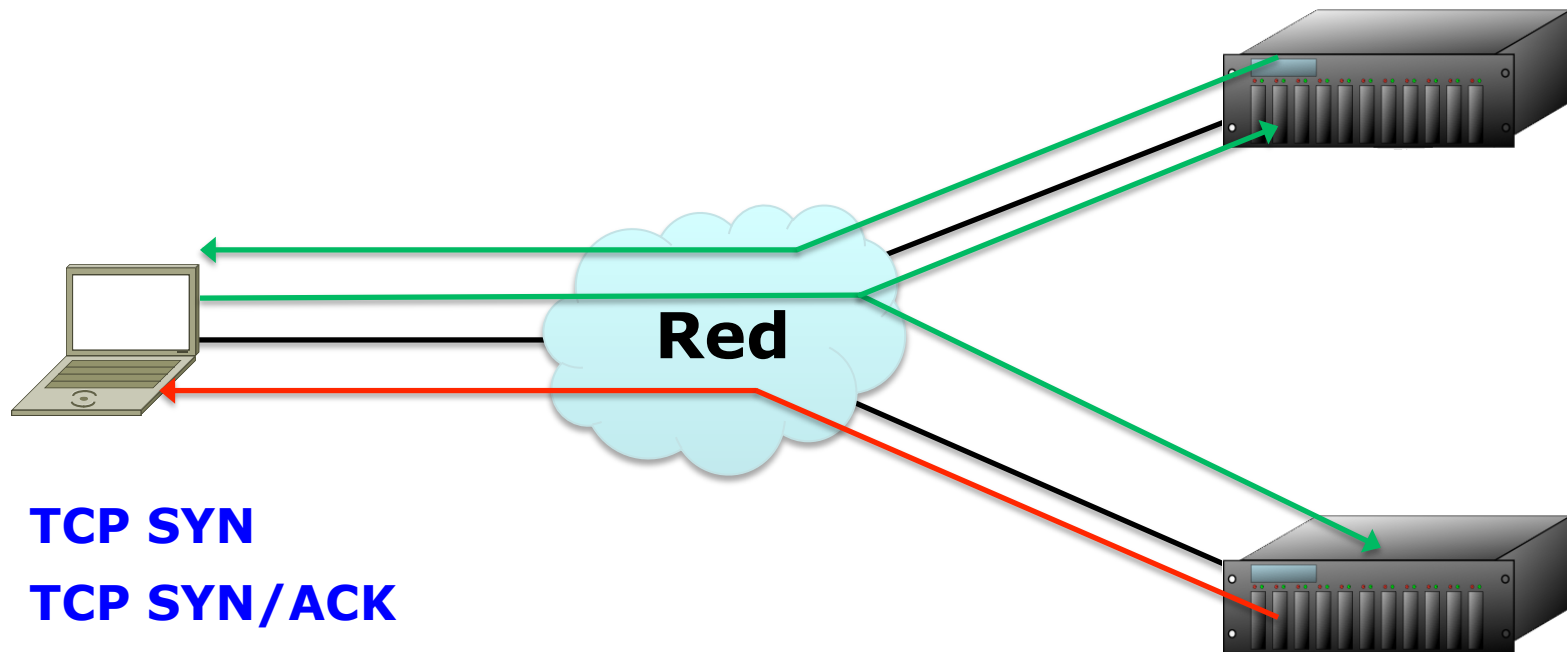


Impacto en Protocolos: TCP

- Sólo funciona cuando el camino en la red es estable
 - **Nunca es cierto en el largo plazo**, pero usualmente es *suficientemente* cierto para períodos cortos
- **El usuario tiene que llegar al mismo destino Anycast durante toda la conexión**
 - Un sólo paquete llegando al destino equivocado provoca un *TCP Reset* que generalmente rompe la conexión



Impacto en Protocolos: TCP



TCP SYN

TCP SYN/ACK

TCP ACK/Datos

TCP Reset

Fuentes de Inestabilidad

1. Balanceo de Carga

- Balanceo por-paquete lleva tráfico por diferentes canales – potencialmente a otros servidores
- Balanceo por-flujo típicamente emplea una 5-tupla para lograr estabilidad en muchas topologías – pero existen casos patológicos

2. Pérdida de Rutas

- Fallas de {Circuito,Enrutador,Servidor}
- Cambios a la configuración: Sesiones que se agregan o quitan, métricas que cambian, ...

3. Dispositivos en el Camino

- Efectos interesantes de los “Optimizadores de Rutas” y balanceadores de carga sobre los flujos de paquetes

Impacto en Protocolos: TCP

- ¿Qué significa para la operación?
 - La ubicación de los servidores Anycast es importante y depende de la topología y configuración
 - Funciona bien para conexiones cortas, cuando se despliega correctamente
 - A mayor duración de la conexión, mayor riesgo de falla
- No es sólo su red, es **cada red que recorren los paquetes** hasta el servidor Anycast
- Evite Anycast de TCP cuando existan (¿buenas?) alternativas

Explorando

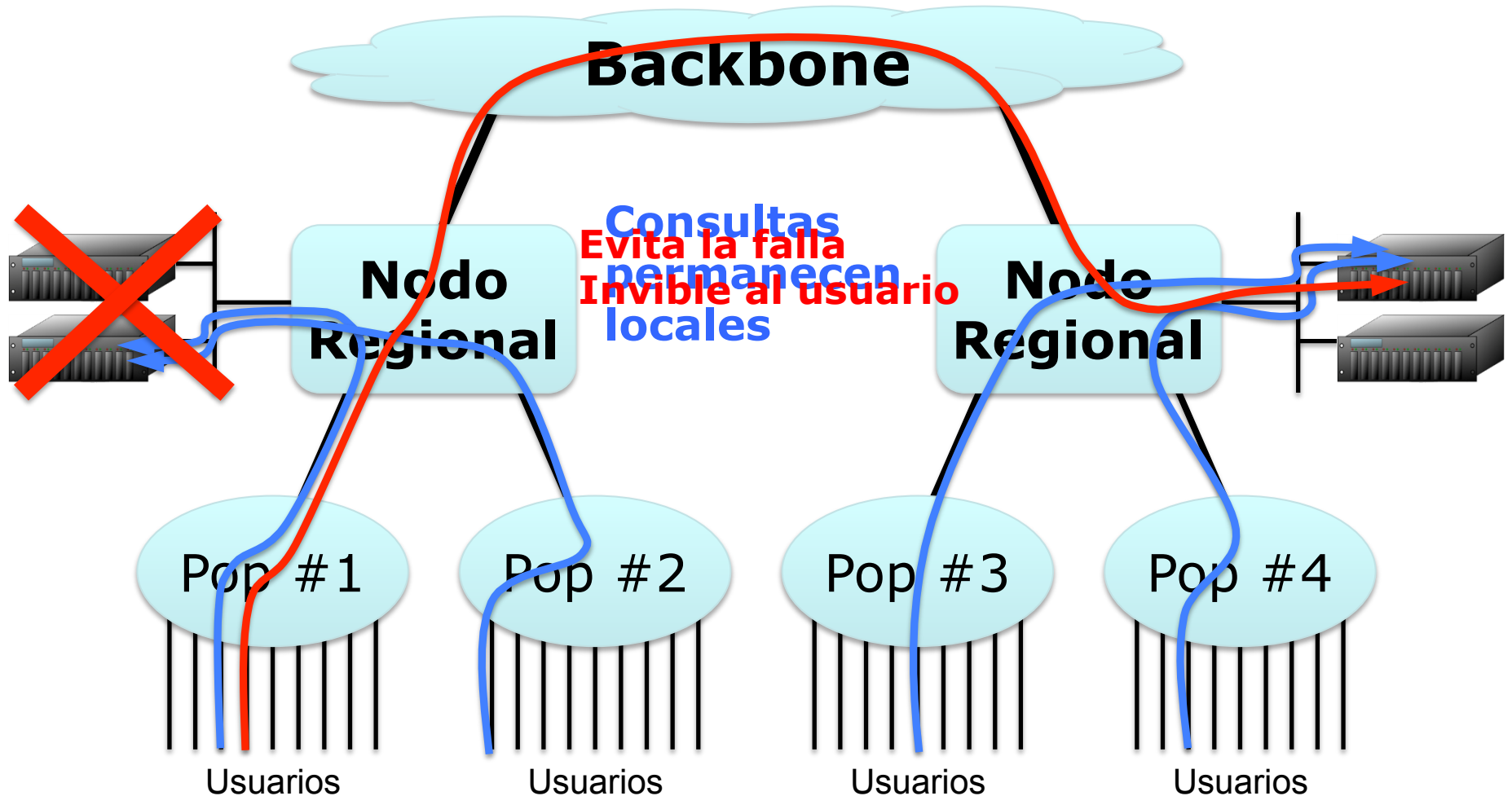
ANYCAST Y DNS



Anycast y DNS

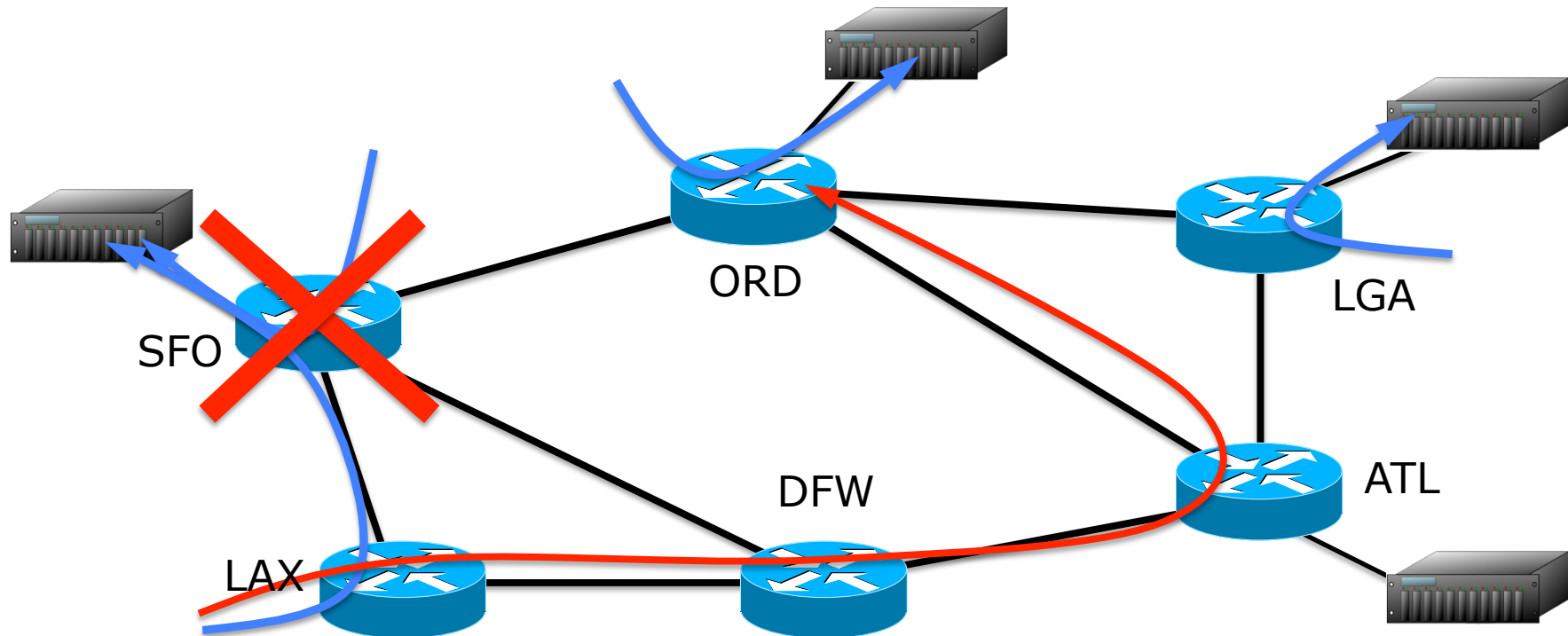
- La mayoría de las consultas comunes son un sólo paquete UDP con 1-3 paquetes de respuesta
- Consultas TCP muy cortas
 - Usuario → Servidor: SYN, ACK c/consulta, ACK/FIN
 - Servidor → Usuario: SYN/ACK, ACK c/datos, ACK/FIN
 - A veces hay un paquete de datos adicional
 - Algunas implementaciones reciben los datos aunque se pierda el FIN
- Las transferencias de zona son consultas TCP longevas
 - Largo depende dell tamaño de la zona
 - Algunas zonas no permiten mitigar el problema

Resolución para Suscriptores



DNS y Anycast

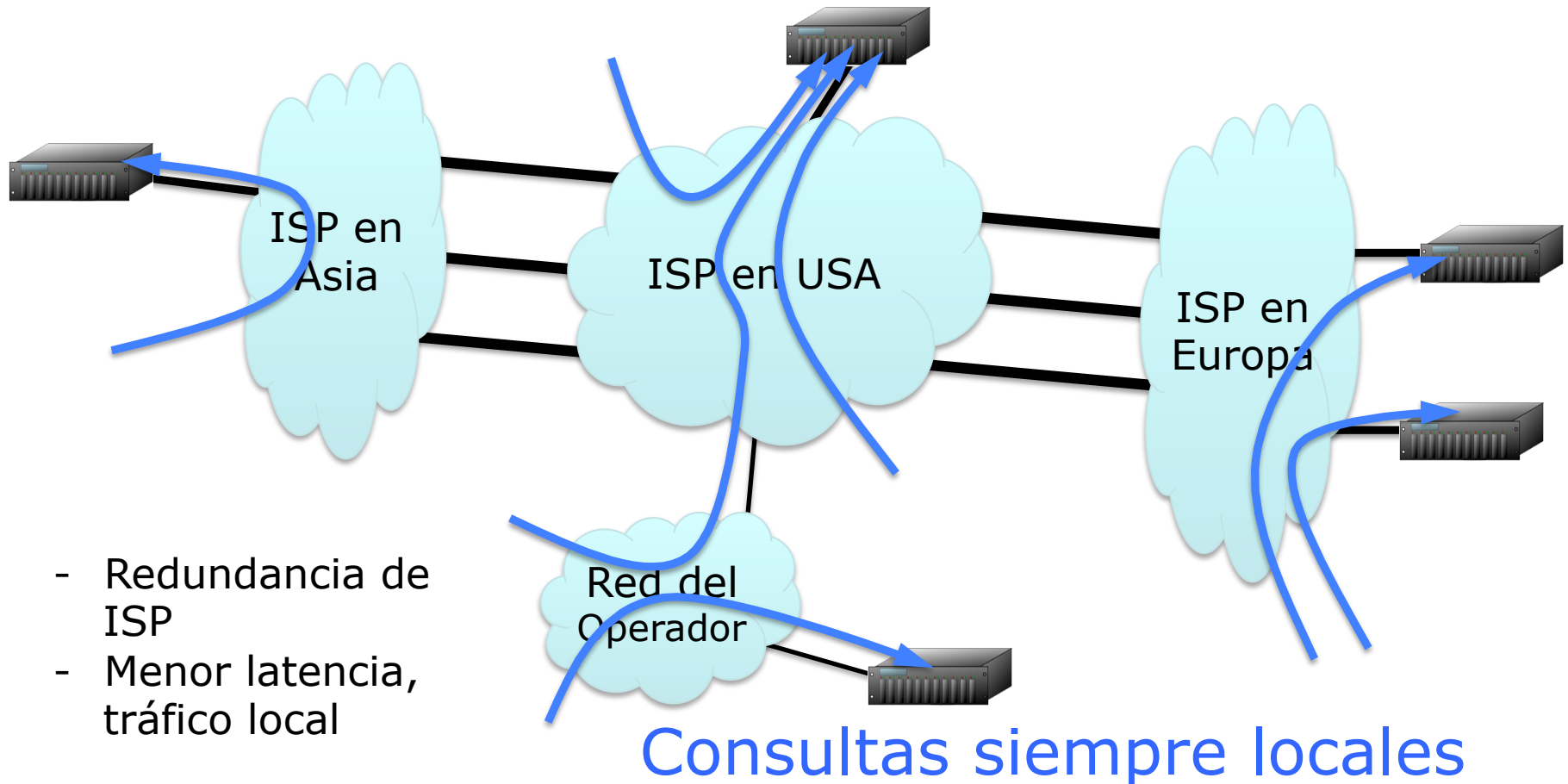
- Servidores de autoridad diseminados ofrecen redundancia, balanceo de carga y mantenimiento sencillo



Falla de un POP, **Consultas locales** se mantiene el servicio

DNS y Anycast

- Servidores de autoridad en múltiples redes



DNS y Anycast: Nivel Avanzado

- Contenido inconsistente
 - Uno de los *ingredientes secretos* que usan algunos CDNs
 - Cada servidor Anycast tiene datos ligeramente distintos y devuelve respuestas que dirigen convenientemente a los usuarios
 - El servidor Anycast sólo sabe sobre el *resolver* al que el usuario hizo la consulta
 - Eso puede ser suficiente en muchos casos
- Usar mecanismos de enrutamiento para dirigir el tráfico de maneras interesantes
 - Usar múltiples super/subredes
 - Métricas que cambian dinámicamente
 - Agregar y quitar rutas basándose en el estado de los servidores

Compartir

EXPERIENCIA OPERATIVA DE ISC



SNS@ISC

- Producto de hospedaje autoritativo de ISC
- Presencia en 3 ISPs globales diferentes
 - Cogent, Hurricane Electric, Tata Communications
- Anycast *dentro* de cada ISP
 - Se emplea espacio IP de cada ISP dentro de su propia red
 - Mínimo de 3 localidades en la red de cada ISP
- Incluyendo 3 registros NS en una zona, el servicio está disponible en 9 ubicaciones en el mundo, via 3 ISPs diferentes
- Cada NS bajo un TLD diferente

SNS@ISC



F-Root

- Tres niveles de Anycast

- LAN Local

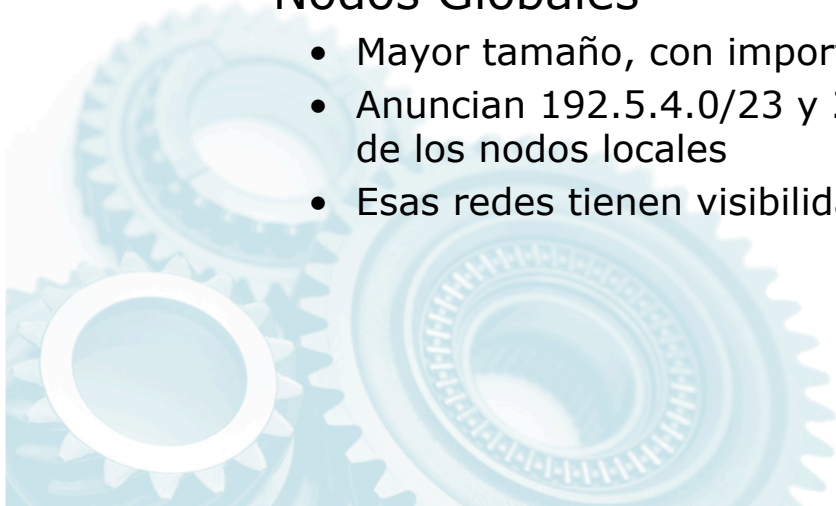
- Cada nodo tiene un mínimo de 2 servidores en la red local para redundancia. Más cuando se requiere

- Nodos Locales

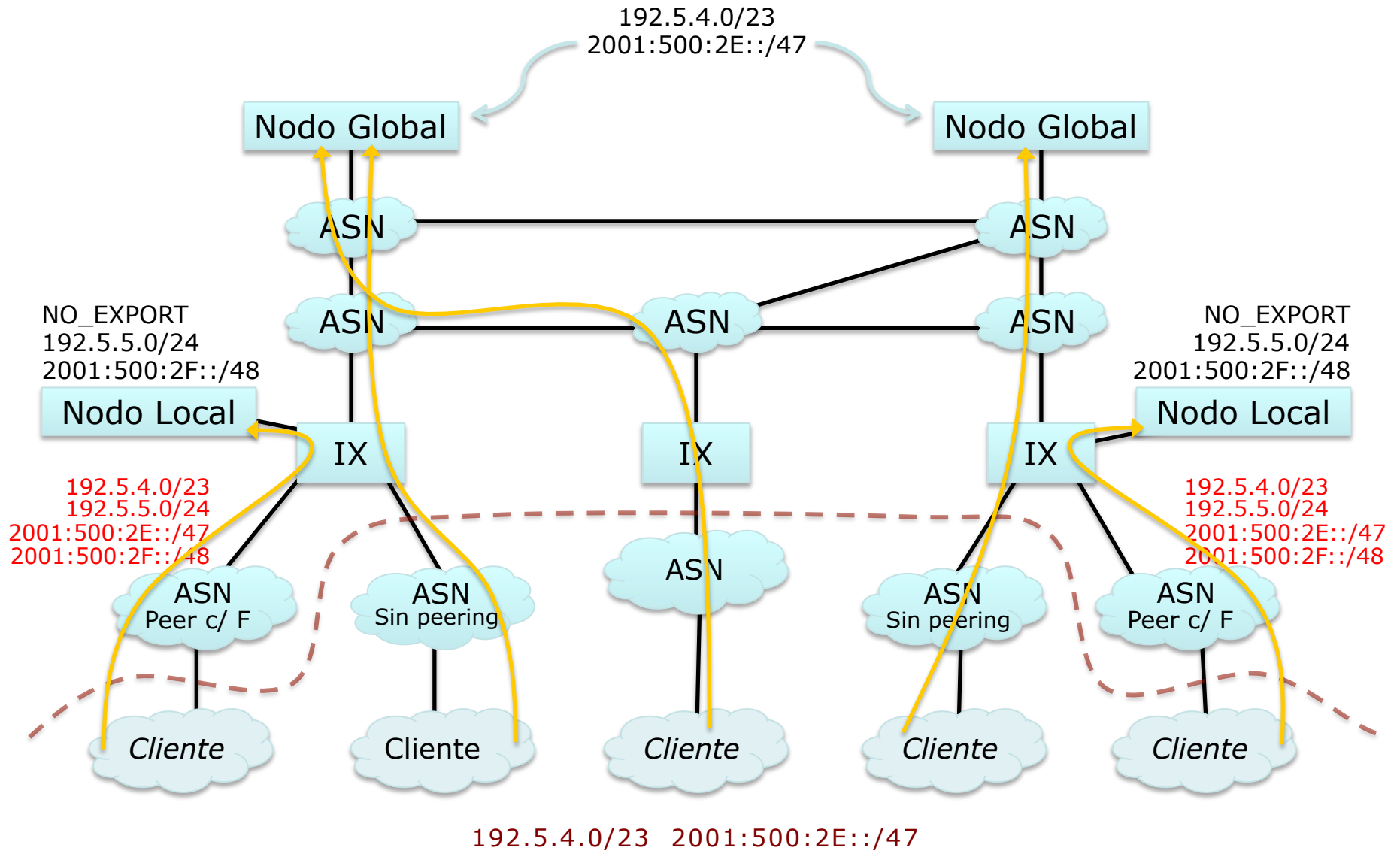
- Típicamente se ubican en un punto de intercambio o dentro de una red de un ISP
 - Anuncia 192.5.5.0/24 y 2001:500:2f::/48 con NO_EXPORT
 - Gracias a NO_EXPORT, esas rutas son visibles selectivamente

- Nodos Globales

- Mayor tamaño, con importante capacidad de tránsito
 - Anuncian 192.5.4.0/23 y 2001:500:2e::/47, super-redes para los anuncios de los nodos locales
 - Esas redes tienen visibilidad global, a todos los usuarios finales



F-Root



F-Root

- ¿Por qué 3 niveles?

- Es importante mantener la localidad del tráfico
 - Los nodos locales pueden servir áreas con poco ancho de banda, p.ej. Cuando se sirven con enlaces satelitales, por lo que no deben *atraer* consultas de otros sitios
 - Incentivo para que los ISPs locales establezcan *peering* con la instancia F-Root local.
- Diversidad en el Ecosistema de Servidores Raíz
 - Los operadores de la raíz creen que tener diferentes grupos desplegando diferentes modelos resulta en un servicio más efectivo a las diversas comunidades de usuarios y ofrece una superficie más resistente al ataque
 - ¡Nadie más usa este método!

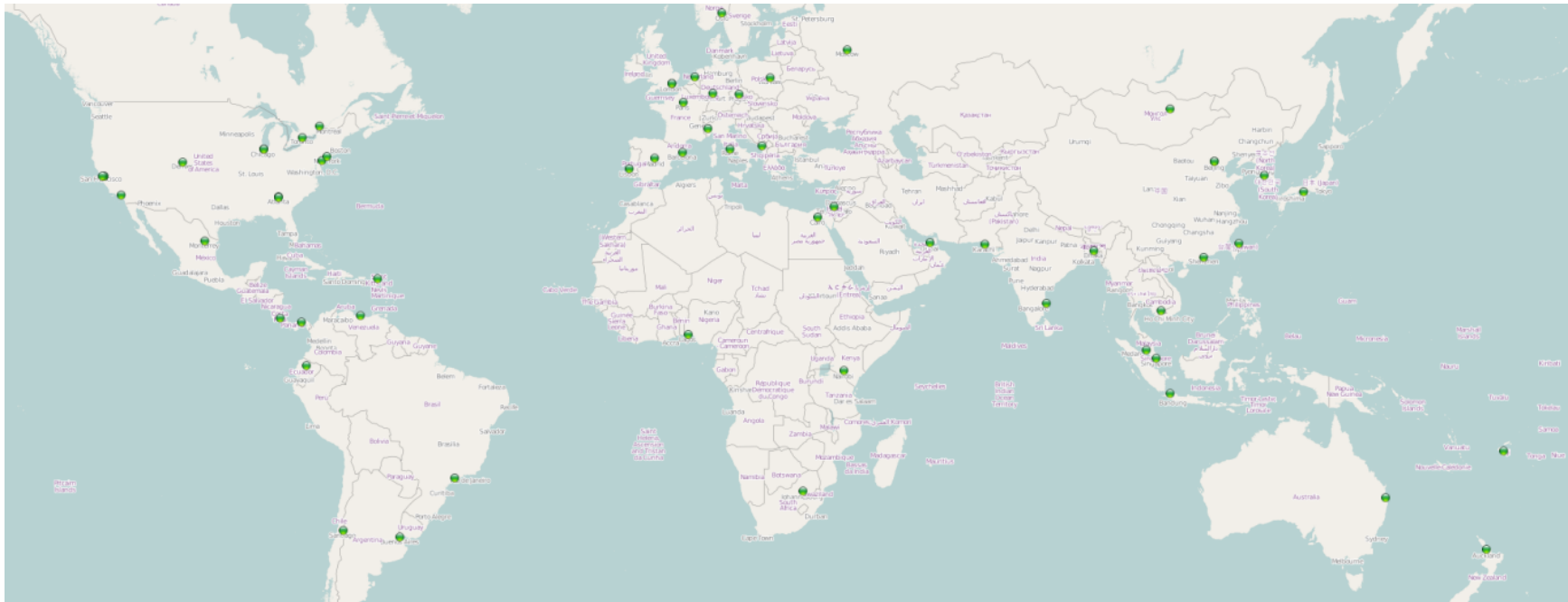
- Se crea confusión

- Algunos ISPs creen – erróneamente – que el NO_EXPORT en las rutas locales impide que sus clientes. Olvidan la super-red.

F-Root

- La transferencia de zonas no está oficialmente soportada, pero se permite
 - Si las conexiones TCP longevas fallan, ISC no lo considera un problema
- Antes de IPv6 y DNSSEC las conexiones TCP eran muy raras
 - 0.00% antes de DNSSEC
 - 0.2% - 0.4% con DNSSEC
 - La mayoría de las implementaciones de DNS manejan el caso de un servidor que no responde intentando con otro servidor
- Considere combinar servidores de autoridad Anycast y Unicast

F-Root



Para terminar

ANYCAST



Resumen

- Anycast es un esquema de enrutamiento útil al desplegar algunas aplicaciones
- Implicaciones a nivel de protocolo que deben ser consideradas en el diseño del despliegue Anycast
- Normalmente DNS está bien adaptado para el uso con Anycast, donde es muy popular
- Muchos otros ya lo están haciendo - ¡No tema!

Conocer más

EVENTOS DE ISC



Eventos y Cursos

www.isc.org/webinars

- Cyber Crime Remediation
 - 22 May 2012
- IPv6 Lessons Learned
 - 12 June 2012

www.isc.org/support/training

- 3-Day IPv6 Fundamentals
 - 4-6 June 2012, Amsterdam
- 2-Day DHCP Workshop
 - 7-8 June 2012, Amsterdam
- 2-Day Intro DNS & BIND
 - 18-19 June 2012, Virginia
- 5-Day Adv DNS & BIND
 - 18-22 June 2012, Virginia
- 2-Day Intro DNS & BIND
 - 2-3 July 2012, Amsterdam
- 5-Day Adv DNS & BIND
 - 2-6 July 2012, Amsterdam



Oferta Especial

Descuento del 18% en cualquier sesión de entrenamiento hasta el 30 de septiembre 2012!

¡Nuestro agradecimiento por acompañarnos!

El cupón electrónico acompañará el mensaje de cierre de este evento.

www.isc.org/support/training



Saber más

PREGUNTAS Y RESPUESTAS

Sigamos en contacto



www.facebook.com/InternetSystemsConsortium



www.linkedin.com/company/internet-systems-consortium



www.twitter.com/ISCdotORG

