Diseño e implementación de una solución de administración de tráfico de red basada en DNS y chequeos de disponibilidad

Tesis para obtener el grado de Magister en Redes de Datos

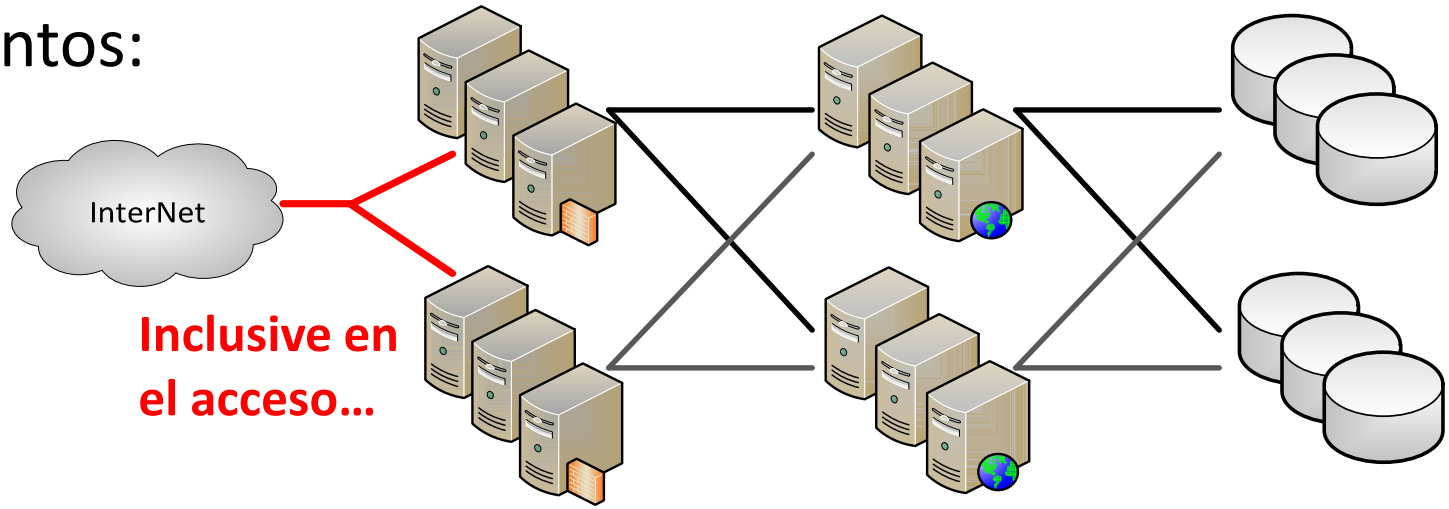
Autor: Nicolás del Río
Directora: Mg. Lía Molinari
Codirector: Ing. Luis Marrone

AGENDA

- Motivación
- Objetivo propuesto
- Tecnologías utilizadas:
  - BGP
  - DNS
  - Chequeos de Disponibilidad
- La herramienta Traffic Manager
- Una pequeña demostración
- Conclusiones finales

MOTIVACIÓN

- Los servicios prestados a través de la red han crecido exponencialmente en los últimos años
- La infraestructura de hardware y comunicaciones, generalmente no acompaña al crecimiento
- Brindar servicios con un alto nivel de disponibilidad es un gran desafío para los administradores
- La disponibilidad debe considerarse en todos los puntos:



OBJETIVO PROPUESTO

- ***Analizar*** e ***implementar*** una solución de administración de tráfico de red basada en el protocolo DNS y chequeos de disponibilidad
- ***Proveer*** una herramienta que brinde alta disponibilidad de red en el acceso a los servicios
- ***Demostrar*** la conveniencia en la utilización de la herramienta, frente a otras soluciones de alta disponibilidad en el acceso

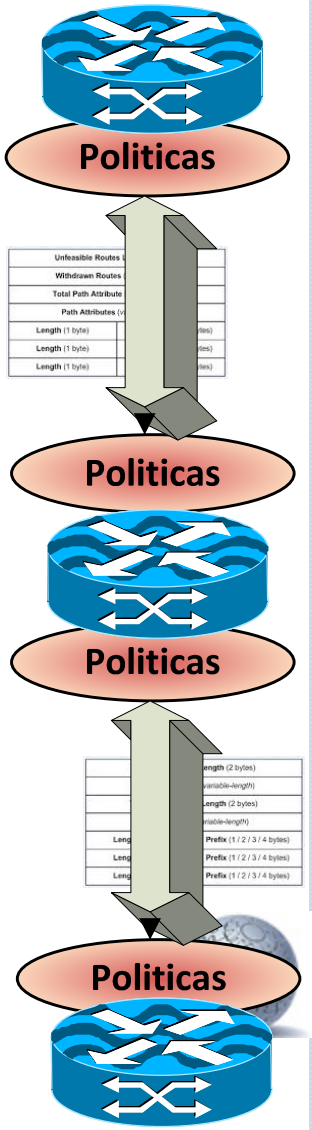
Traffic Manager

TECNOLOGÍAS

BGP

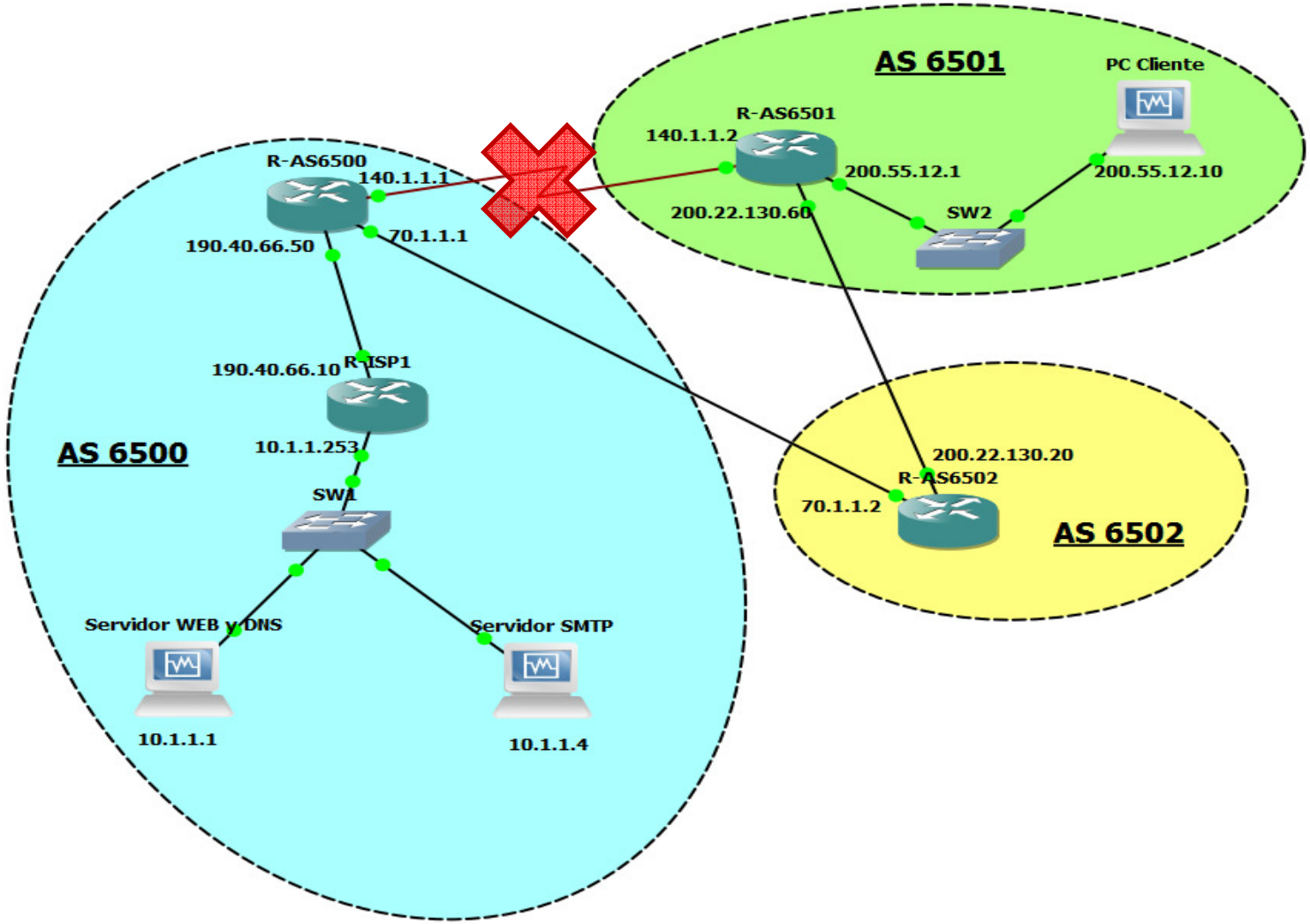
UNA APROXIMACIÓN A LA SOLUCIÓN

- **B**order **G**ateway **P**rotocol es un protocolo de ruteo que permite el intercambio de información de ruteo entre organizaciones (AS)
- IPv4: RFC 4271 - IPv6: RFC 2460
- Opera en la capa 3 de modelo OSI
- Permite que un destino de la red pueda ser alcanzado por más de 1 camino
- Es un protocolo muy potente gracias a la posibilidad de implementar políticas



BGP

UN EJEMPLO PARA COMPRENDER LA TECNOLOGÍA



La información contenida en las gráficas es a modo ilustrativo y no representa información real

BGP

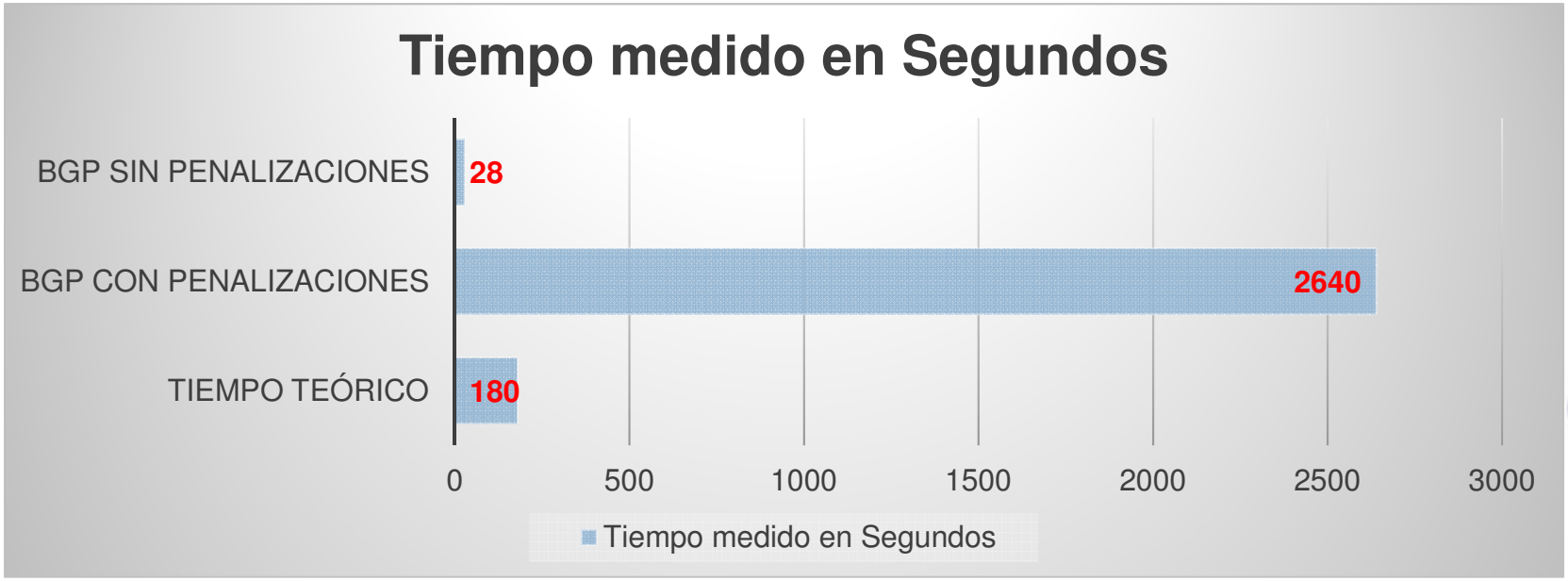
Convergencia y Tiempos

- Cuando se produce un cambio en la red, el router informa a sus vecinos para que actualicen su información
- Sucesivas notificaciones de cambios pueden sufrir penalizaciones
- La técnica de penalizaciones permite limitar la utilización de recursos por parte de routers que realicen sucesivas modificaciones a la tabla de ruteo

BGP

Convergencia y Tiempos

- En el capítulo 6 se presentan 2 escenarios de pruebas basados en la topología anterior, con sus tiempos de convergencia:
  - 6.3: Tiempo de convergencia sin penalizaciones: 28 segundos
  - 6.4: Tiempo de convergencia con penalizaciones: 44 minutos



BGP

Registro de IPv4 y número de Sistema Autónomo

- Para Latinoamérica y Caribe, la solicitud debe realizarse ante Lacnic.
- Requisitos para solicitar un bloque IPv4:
  - Informar planificación de uso y ocupación
  - Solicitar un bloque de direcciones IPv6 si la organización aún no lo posee
  - Abonar costo de registro y mantenimiento anual (U$D 2500 y U$D 600 respectivamente)
- Requisitos para solicitar un número de sistema autónomo:
  - Informar una política de ruteo que difiera de la de su proveedor
  - Demostrar que a red contará con más de una conexión independiente a InterNet (esquema multiproveedor)
  - Abonar costo de registro de U$D 1000

Requerimientos y costos vigentes a Febrero de 2015

BGP

CONCLUSIONES

- Se trata de un protocolo maduro y estable que provee alta disponibilidad en el acceso
- La posibilidad de implementar políticas, permite manipular el flujo natural de los datos
- Se adapta simplemente a los cambios de red con tiempos de convergencia que van desde 28 segundos a 44 minutos
- Los requerimientos para trabajar con BGP hacen que el mismo no sea accesible para cualquier usuario (requisitos formales / costo)
- Su uso y configuración es complejo

DNS y Chequeos de disponibilidad

Otra solución

BGP

DNS

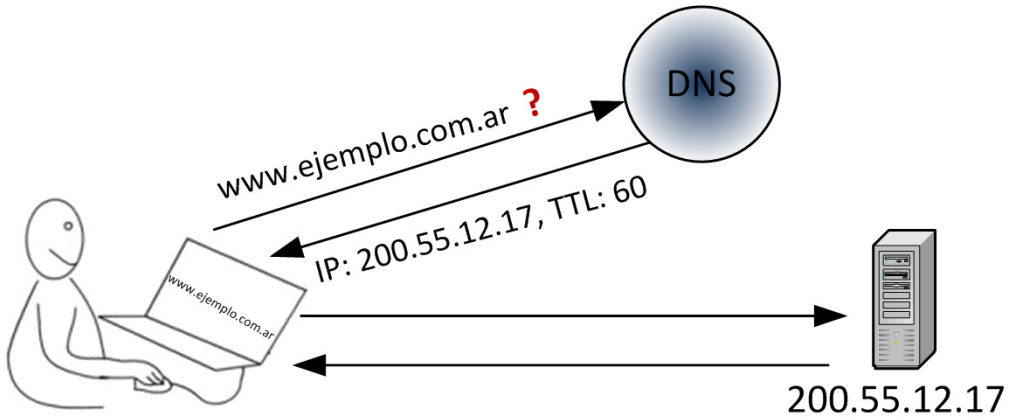
+

Chequeos de Disponibilidad

DNS

Comprendiendo el protocolo para encontrar similitudes

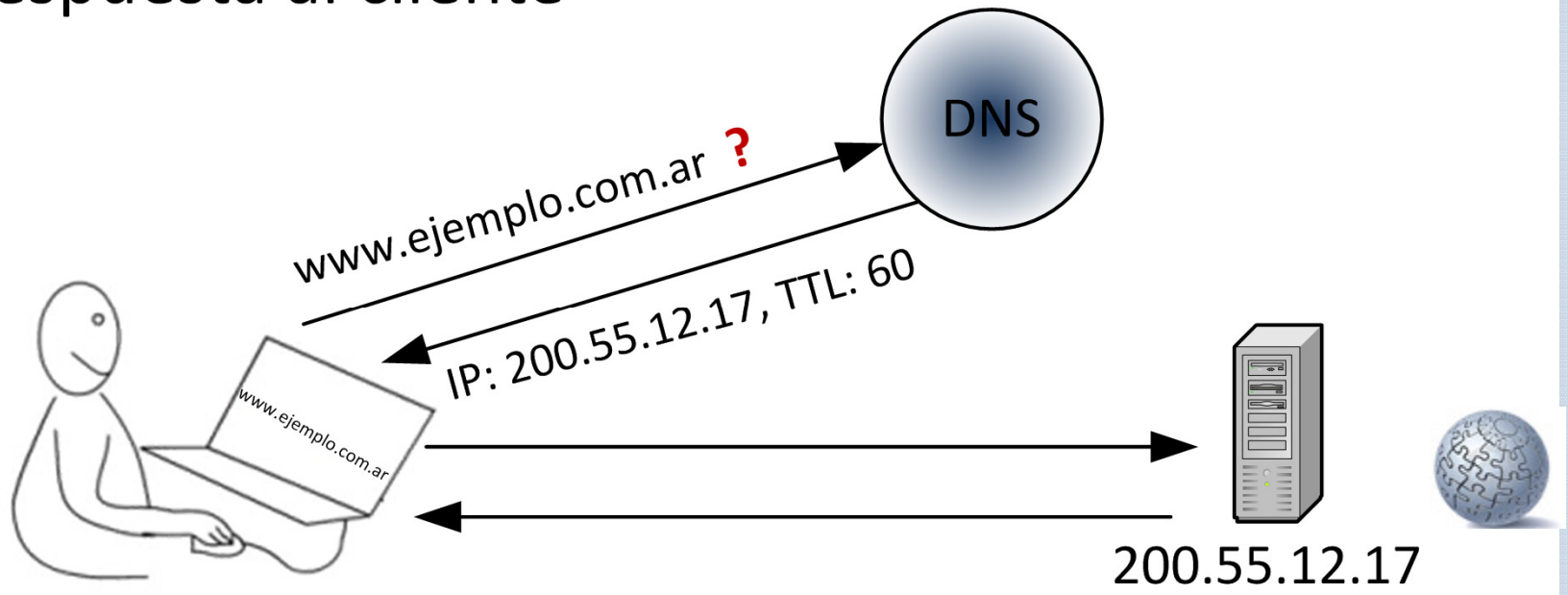
- **D**omain **N**ame **S**ystem es un sistema distribuido de nomenclatura jerárquica que asocia nombres con direcciones IP y viceversa, entre otras cosas
- Se encuentra estandarizado en la RFC 1034 y 1035
- Opera en la capa 7 de modelo OSI
- Cuando un cliente requiere el acceso a un recurso, envía una solicitud de resolución de nombre a IP a su servidor DNS



DNS (CONT.)

COMPRENDIENDO EL PROTOCOLO PARA ENCONTRAR SIMILITUDES

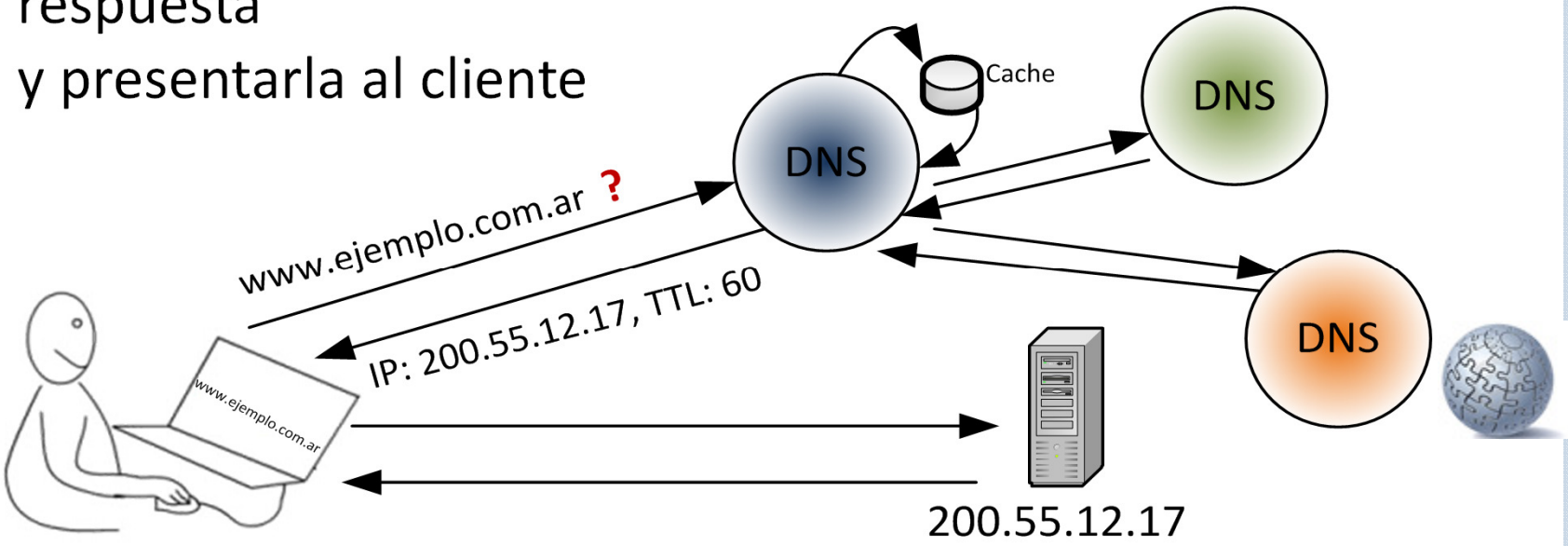
- Cuando un servidor DNS recibe un requerimiento de resolución, verifica su propia base de datos:
  - Si la consulta se corresponde con una porción de la base de datos que él maneja, enviará la respuesta al cliente



DNS (CONT.)

COMPRENDIENDO EL PROTOCOLO PARA ENCONTRAR SIMILITUDES

- Si la consulta no se corresponde con ninguna de las zonas para el cual el servidor es **autoritativo**:
  - Verificará su cache con el fin de determinar si puede enviar directamente la respuesta
  - Si la respuesta no se encuentra en la cache, realizará la consulta a otros servidores DNS con el fin de obtener la respuesta
    y presentarla al cliente



DNS

ALTA DISPONIBILIDAD

- Con el fin de brindar alta disponibilidad en el servicio DNS, se define el esquema de servidor primario y secundario
- El servidor primario contiene la base de datos de resolución donde el administrador modifica los registros
- El servidor secundario almacena una copia sincronizada de la base de datos del servidor primario
- Si los servidores son accesibles a través de vínculos de InterNet de distintos proveedores, ante la caída de uno de ellos, el otro servidor responde los requerimientos

DNS

Registro de un nombre de dominio

- Los registros de dominios .ar se gestionan ante NIC Argentina en http://nic.ar:
  - Se debe crear una cuenta de usuario, y verificar que el dominio se encuentre disponible
  - Se debe delegar el dominio a un conjunto de servidores DNS que responderán por él
  - Abonar costo de registro y renovación anual de $ 220 (pesos Argentinos) para un dominio .com.ar

Requerimientos y costos vigentes a Febrero de 2015

Chequeos de Disponibilidad

- Las infraestructuras de cómputo deben ser monitoreadas en todo momento
- El monitoreo permite detectar fallas y actuar en consecuencia
- Existen distintas aplicaciones de software y mecanismos para realizar los chequeos
- Uno de los más utilizados es Icinga:
  - Es de código abierto
  - Es una bifurcación (fork) de Nagios
  - Permite definir chequeos, períodos de tiempo y manejadores de eventos
  - Provee interfaces ***Rest*** para la integración con otras herramientas

Traffic Manager

Una solución de administración de tráfico basada en DNS y Chequeos de Disponibilidad

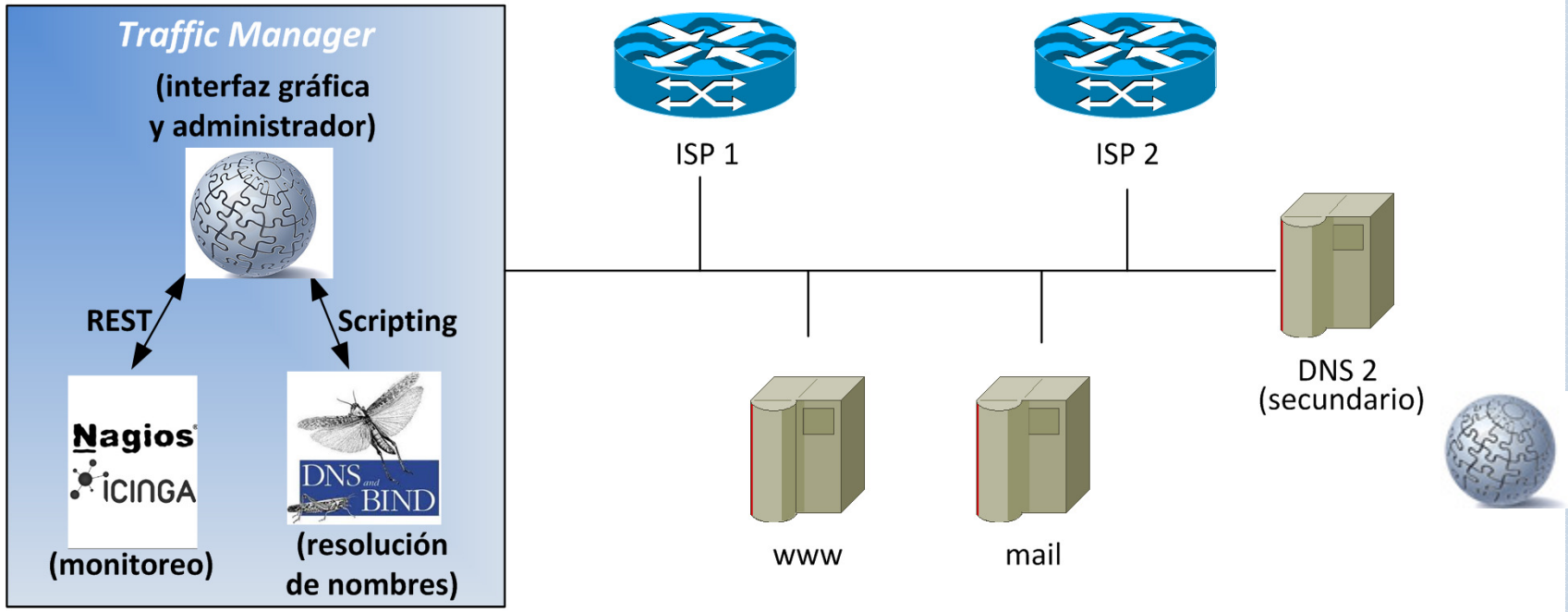
Propuesta de Administración de Tráfico de Red basada en DNS

- El protocolo **DNS** es un pilar básico de la red InterNet:
  - Permite la localización de recursos, a través de resolución de nombre a IP
  - A través de DNS se puede direccionar tráfico, resolviendo los requerimientos con diferentes direcciones dependiendo del estado de la red
- Utilizando de chequeos de disponibilidad se pueden detectar cambios en la red y notificarlos al servicio de resolución para que modifique sus respuestas

En la conjunción DNS + Chequeos de disponibilidad, podemos encontrar cierta semejanza con BGP

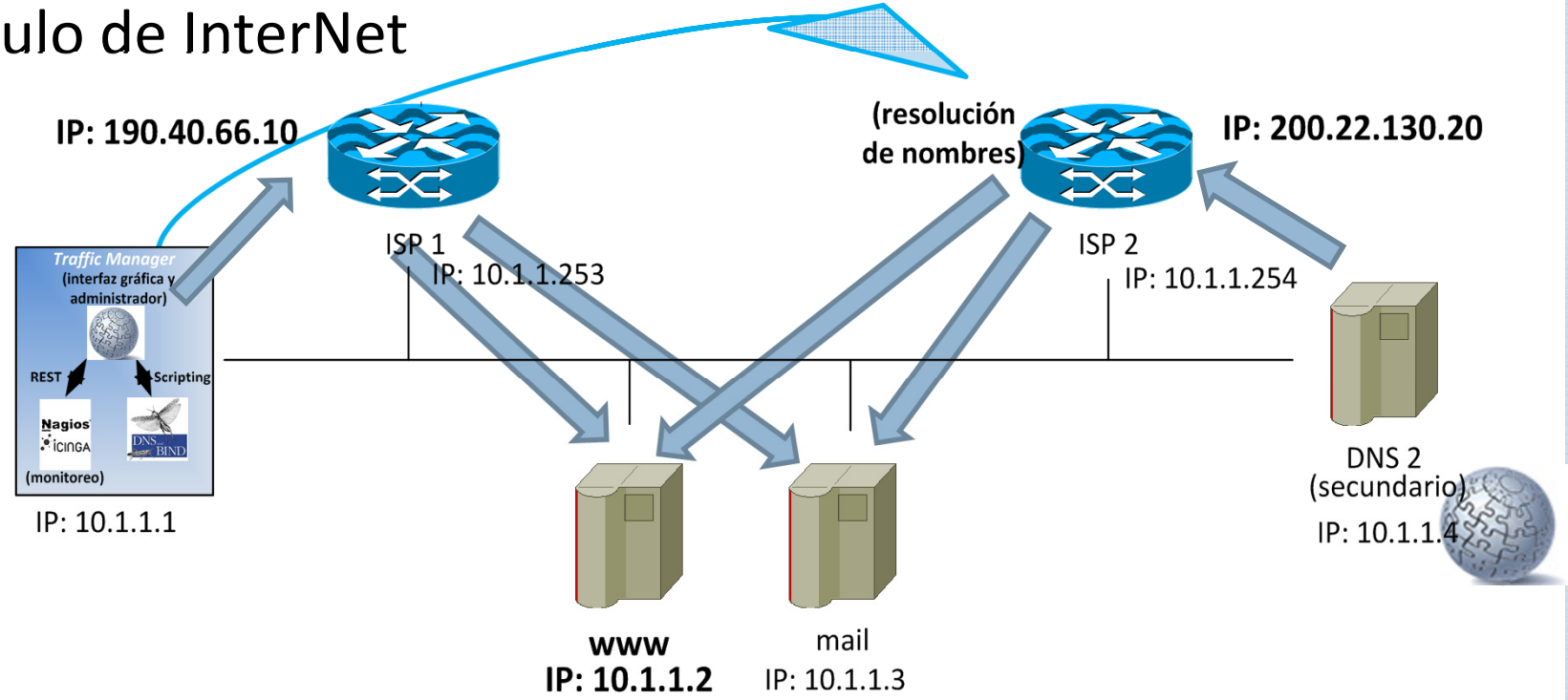
Propuesta de Administración de Tráfico de Red basada en DNS
Arquitectura de la Solución Planteada

- En un esquema multiproveedor, si un servicio es accesible a través de 2 direcciones IP, el DNS podría responder con una u otra indistintamente
- Ante la caída de un vínculo, se puede notificar al DNS para que responda con la IP a través de la cual el servicio sigue siendo accesible



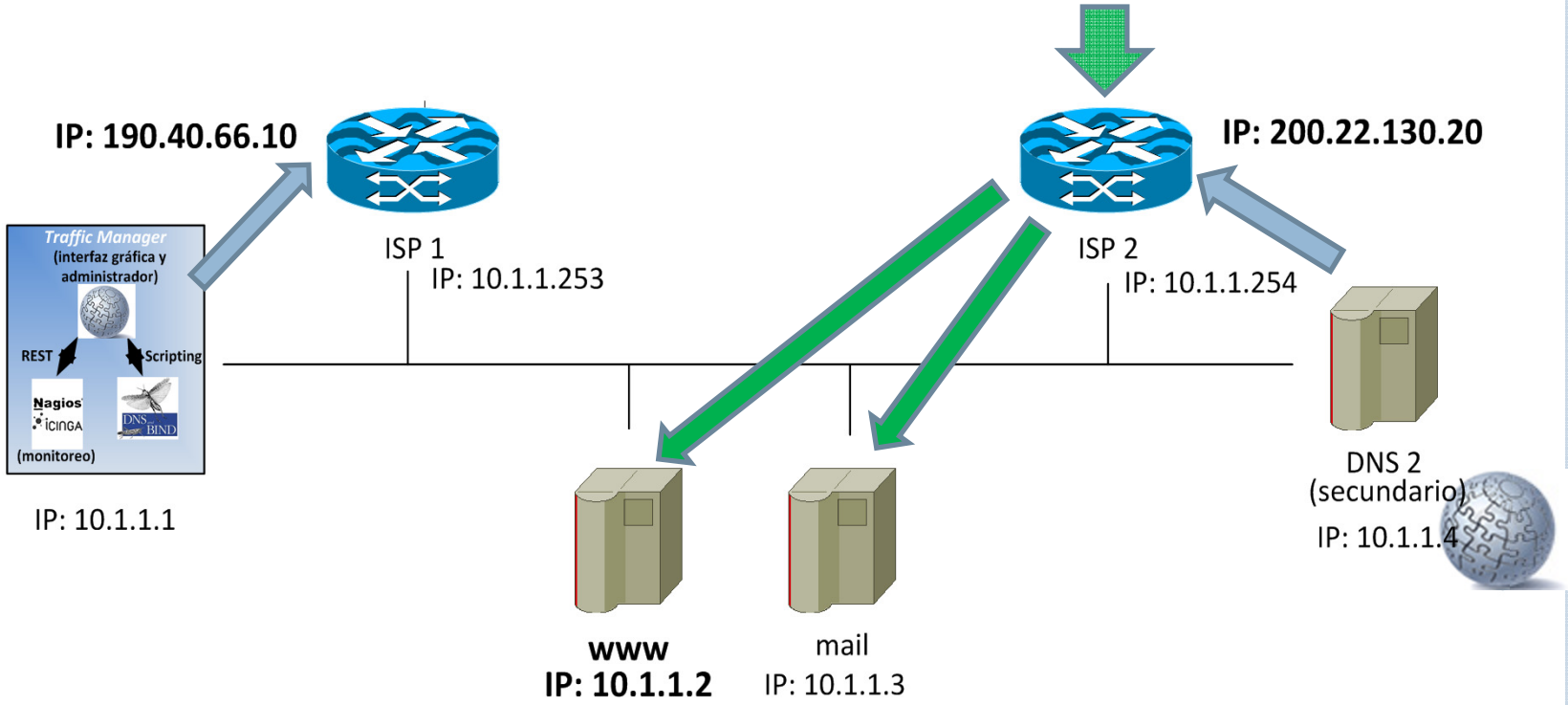
Propuesta de Administración de Tráfico de Red basada en DNS
Arquitectura de la Solución Planteada

- TrafficManager y DNS2 configuran su puerta de enlace a ISP1 e ISP2 respectivamente
- Los routers redirigen DNS, HTTP y SMTP a los servidores
- El servidor TrafficManager chequea regularmente los servicios publicados en el router de ISP2 (HTTP y SMTP) usando el vínculo de InterNet



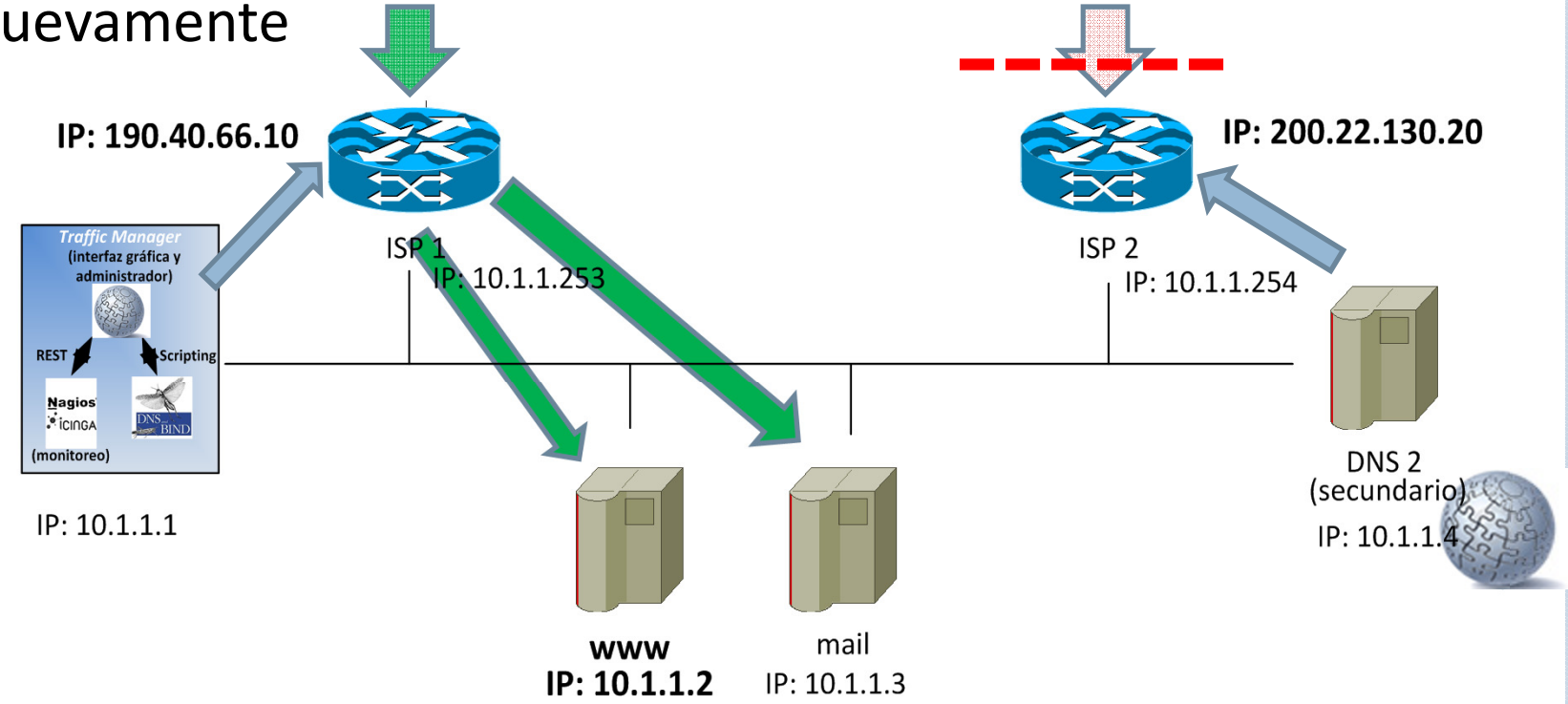
Propuesta de Administración de Tráfico de Red basada en DNS
Arquitectura de la Solución Planteada

- El enlace principal para la publicación de los servicios HTTP y SMTP, es el provisto por ISP2
- El enlace de ISP1 se utiliza para publicar a DNS1 y para realizar chequeos de disponibilidad



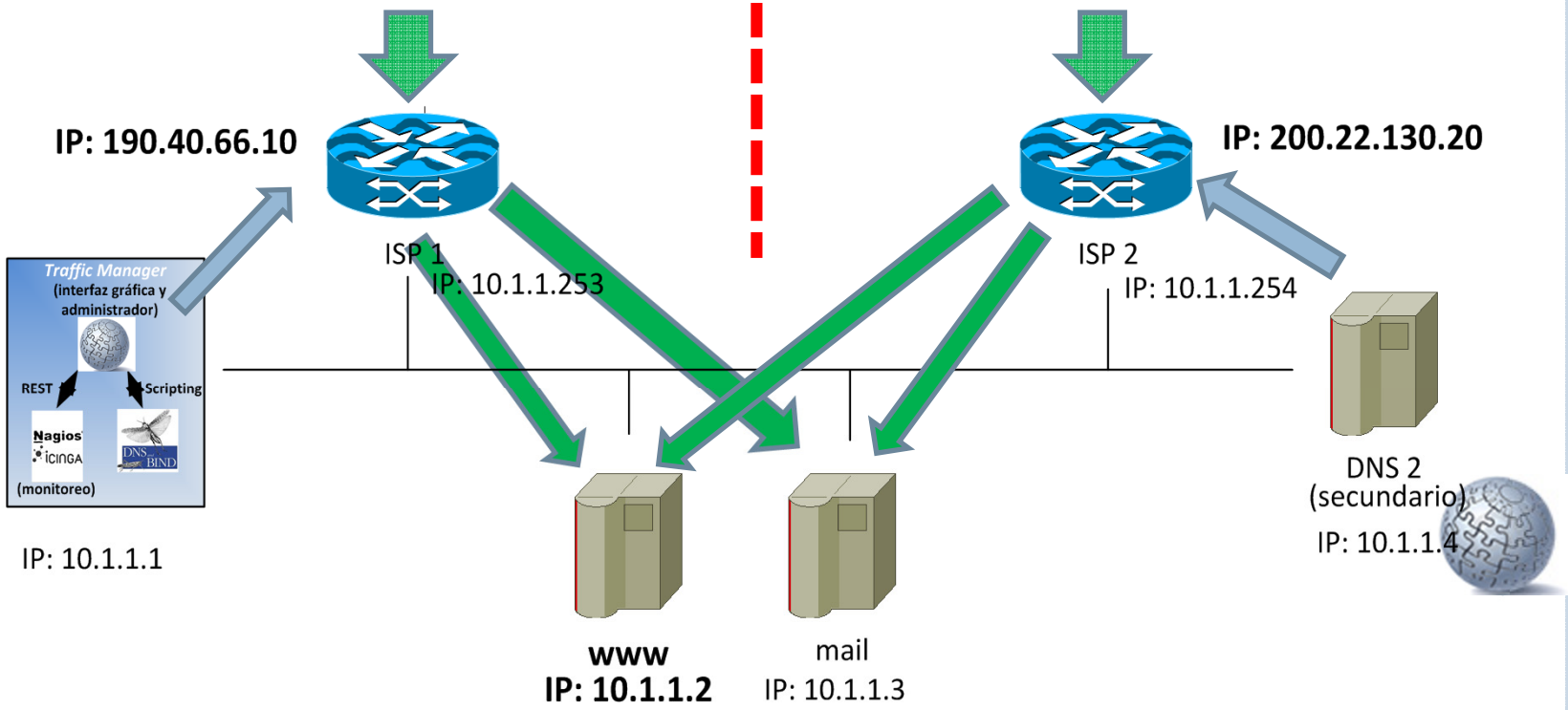
Propuesta de Administración de Tráfico de Red basada en DNS
Escenarios de Fallas que controla la Herramienta

- Ante la caída del vínculo de comunicaciones de ISP2:
  - Se modifican los registros de DNS para cambiar el flujo de los datos
  - Las nuevas resoluciones de DNS responderán con la IP de ISP1
  - Cuando ISP2 esté nuevamente disponible se conmuta nuevamente



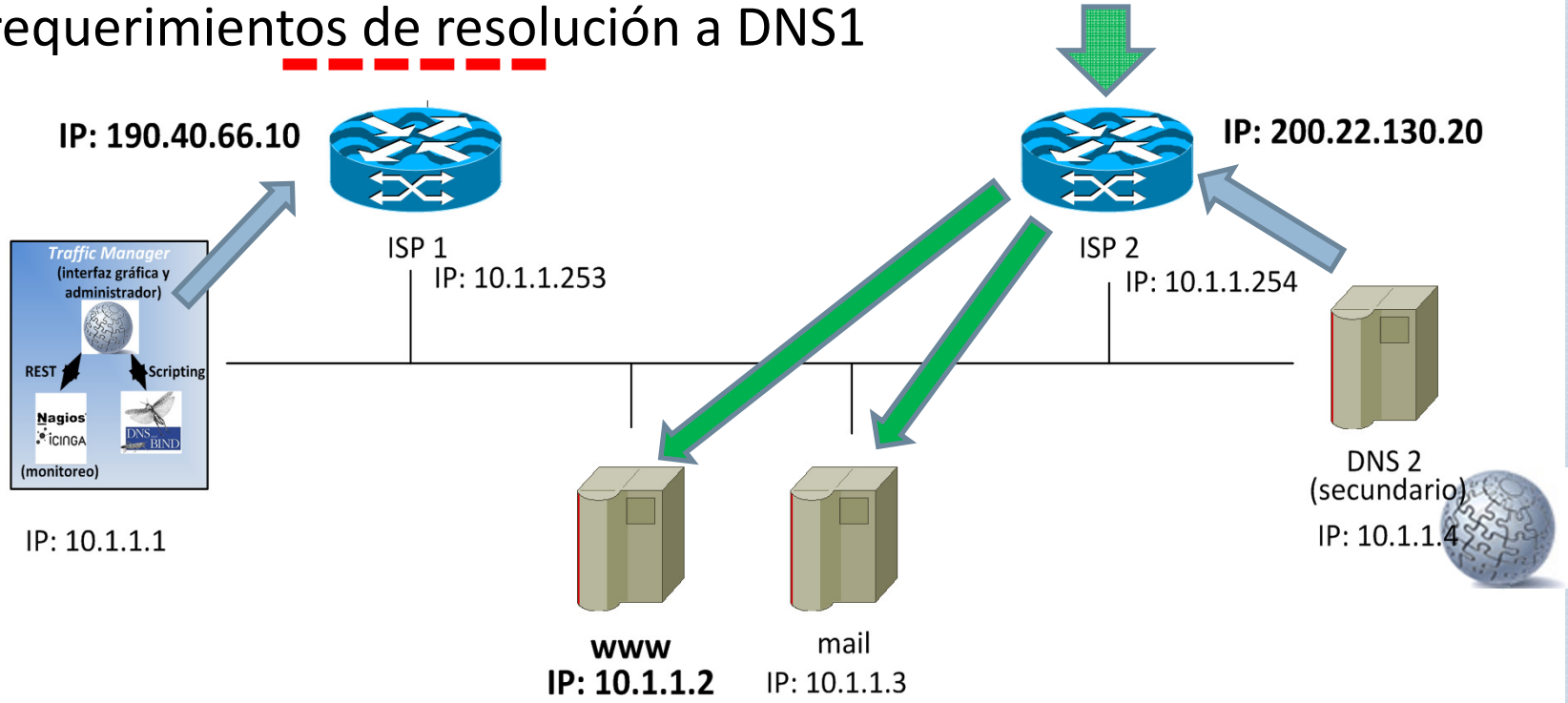
Propuesta de Administración de Tráfico de Red basada en DNS
Escenarios de Fallas que controla la Herramienta

- De existir inconvenientes en un nodo intermedio entre ISP1 e ISP2:
  - Se modifican los registros de DNS para cambiar el flujo de los datos
  - Las actualizaciones de DNS, no llegarán nunca al DNS secundario
  - Quienes consulten a ISP1 obtendrán respuesta con las IP de ISP1
  - Quienes consulten a ISP2 obtendrán respuesta con las IP de ISP2



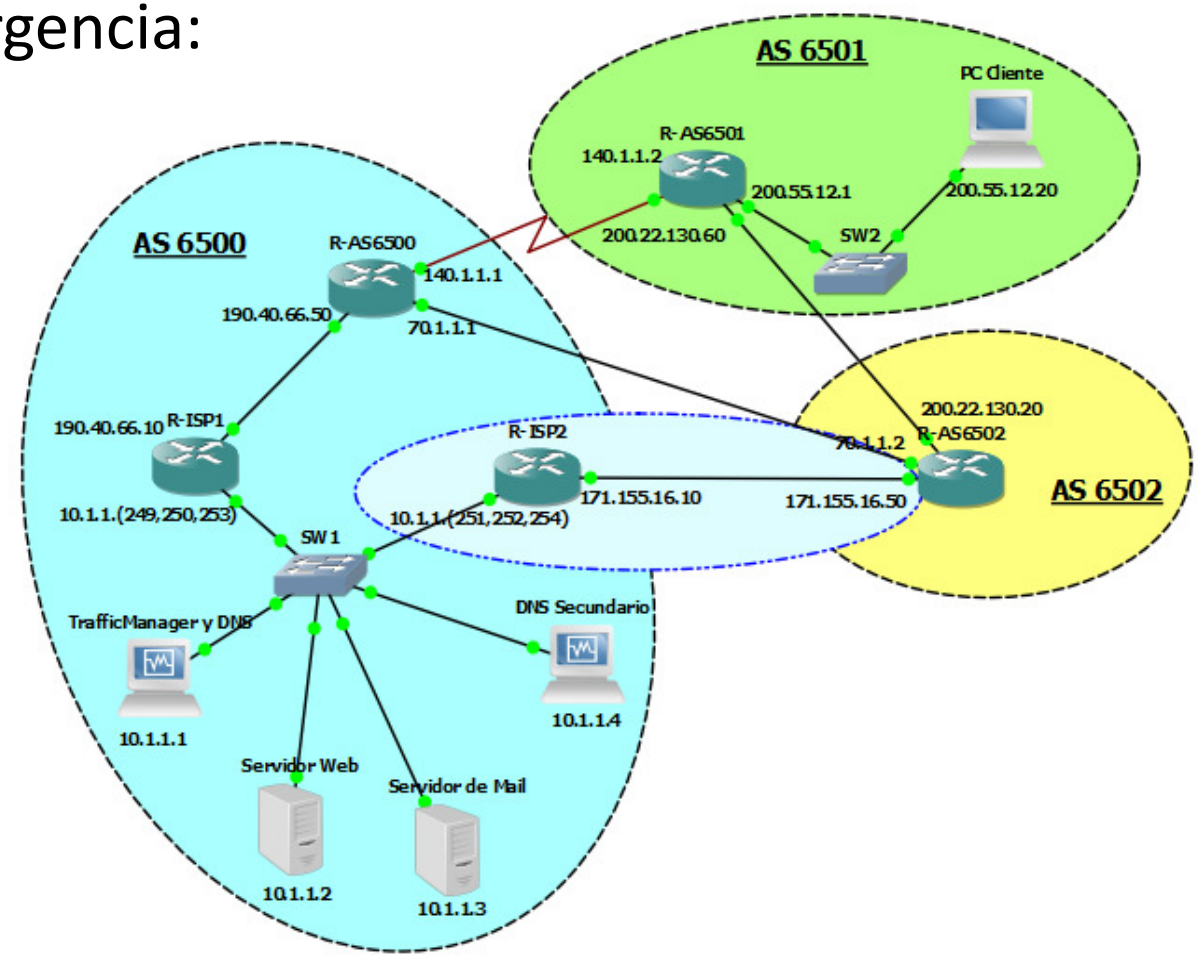
Propuesta de Administración de Tráfico de Red basada en DNS
Escenarios de Fallas que controla la Herramienta

- Ante la caída del vínculo de comunicaciones de ISP1:
  - Se modifican los registros de DNS para cambiar el flujo de los datos
  - Las nuevas resoluciones de DNS responderán con la IP de ISP1
  - Como el vínculo de comunicaciones no está activo, las actualizaciones nunca llegarán a ISP2 y nunca llegarán requerimientos de resolución a DNS1



PROPUESTA DE ADMINISTRACIÓN DE TRÁFICO DE RED BASADA EN DNS
TIEMPOS DE CONVERGENCIA

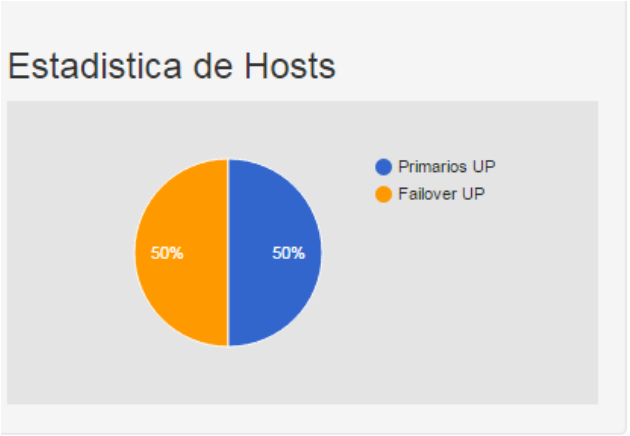
- En el capítulo 6 se presenta 1 escenario de prueba utilizando la herramienta Traffic Manager, con sus tiempos de convergencia:



Propuesta de Administración de Tráfico de Red basada en DNS
Tiempos de Convergencia

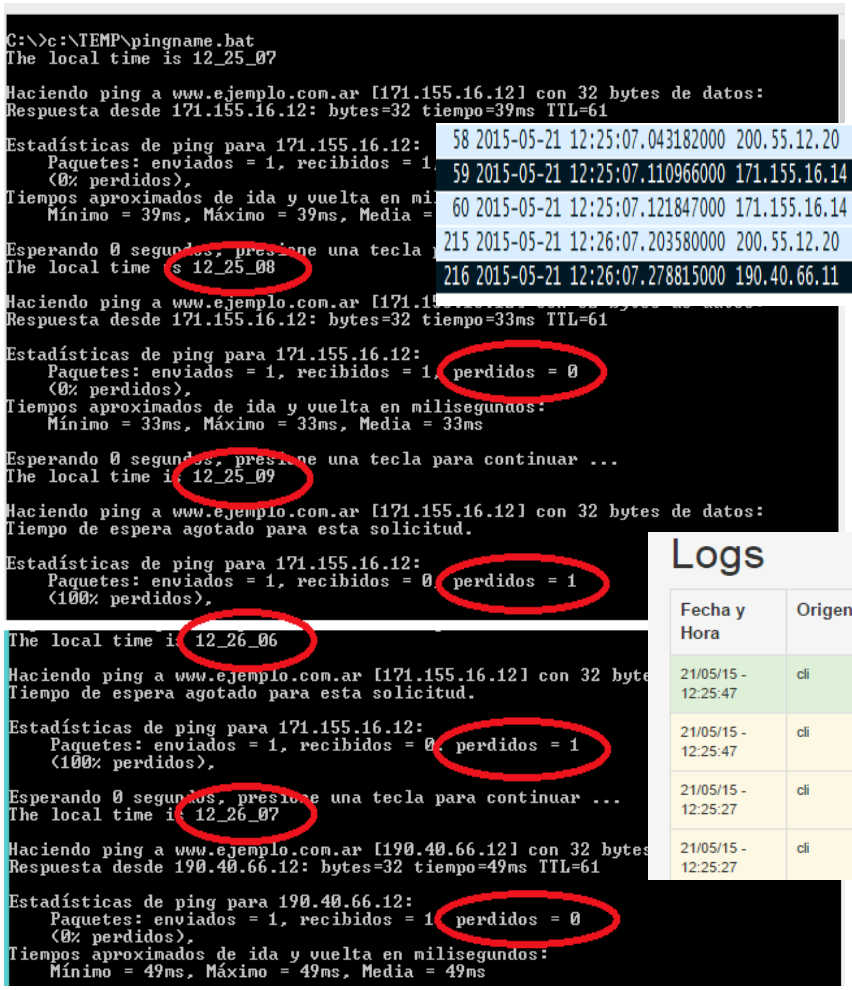
- Se definió el host www en el dominio ejemplo.com.ar con sus correspondientes direcciones IP…

…y se verificó el estado inicial de la red



Propuesta de Administración de Tráfico de Red basada en DNS
Tiempos de Convergencia

- Se generaron cortes de vínculos en la topología y se verificaron los tiempos de convergencia:



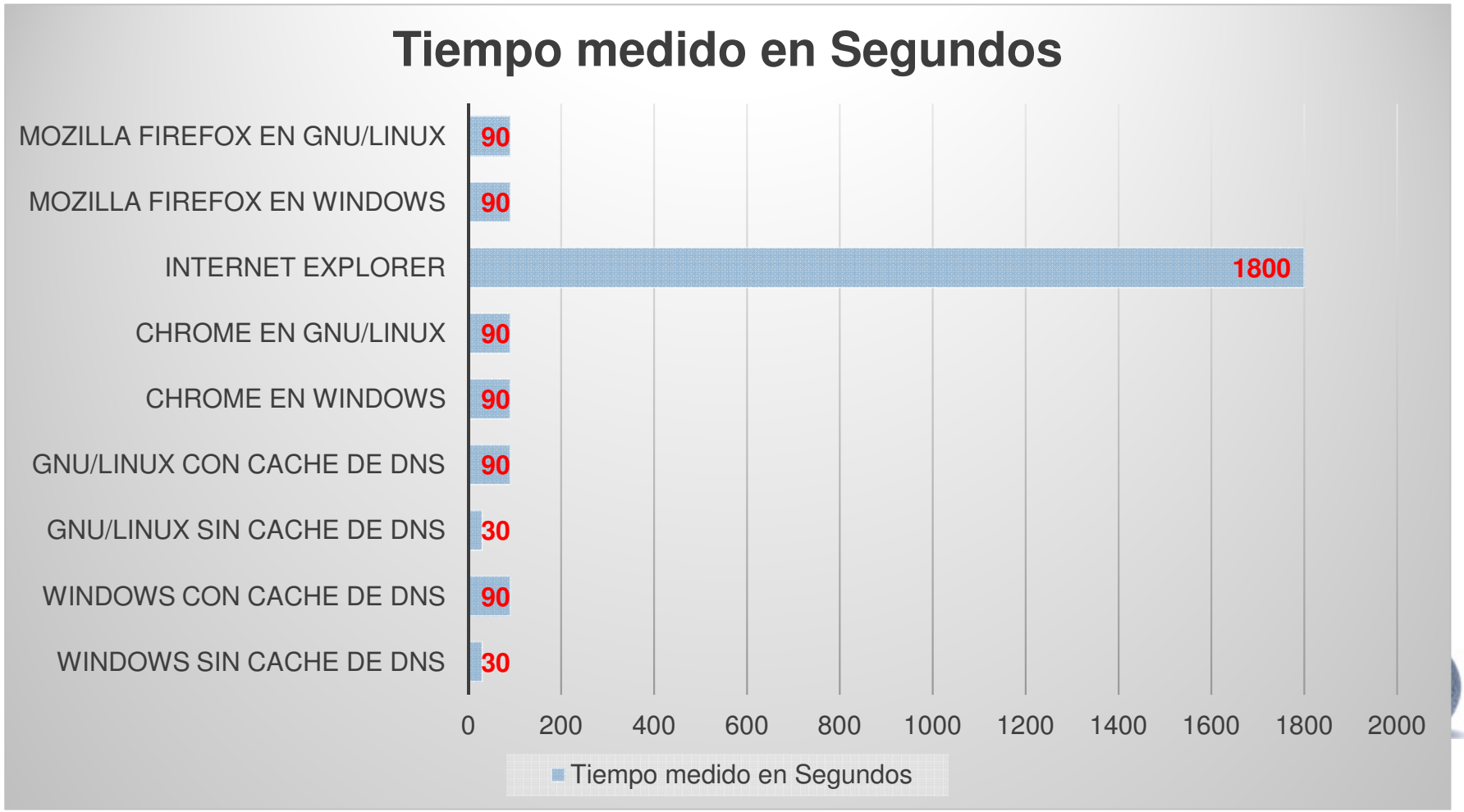
58	2015-05-21 12:25:07.043182000	200.55.12.20	171.155.16.14	DNS	78 Standard query 0x8efd A www.ejemplo.com.ar
59	2015-05-21 12:25:07.110966000	171.155.16.14	200.55.12.20	DNS	164 Standard query response 0x8efd A 171.155.16.12
60	2015-05-21 12:25:07.121847000	171.155.16.14	200.55.12.20	DNS	119 Standard query response 0x5476
215	2015-05-21 12:26:07.203580000	200.55.12.20	190.40.66.11	DNS	78 Standard query 0x2a1c A www.ejemplo.com.ar
216	2015-05-21 12:26:07.278815000	190.40.66.11	200.55.12.20	DNS	164 Standard query response 0x2a1c A 190.40.66.12

Logs

Fecha y Hora	Origen	Modulo	Datos
21/05/15 - 12:25:47	cli	em	Se va a Failover en el HOST www.ejemplo.com.ar. IP de Failover: 190.40.66.12. Estado de HOST principal: DOWN/HARD
21/05/15 - 12:25:47	cli	em	Cambio de Estado en www.ejemplo.com.ar. IP: 171.155.16.12. Failover IP: 190.40.66.12. Estado: DOWN/HARD
21/05/15 - 12:25:27	cli	em	Se detecto una caida en el HOST www.ejemplo.com.ar, del tipo SOFT. Se continuan los chequeos para determinar si es definitiva. No se toma ninguina accion.
21/05/15 - 12:25:27	cli	em	Cambio de Estado en www.ejemplo.com.ar. IP: 171.155.16.12. Failover IP: 190.40.66.12. Estado: DOWN/SOFT

Propuesta de Administración de Tráfico de Red basada en DNS
Tiempos de Convergencia

- Los resultados obtenidos en el peor de los casos fueron:



Propuesta de Administración de Tráfico de Red basada en DNS
Conmutación del Tráfico

- Es disparada a través del servicio de monitoreo al detectar cambios sobre el estado de un servicio
- Es un script PHP que hace interfaz con el servicio DNS usando actualizaciones dinámicas con el comando *nsupdate*:

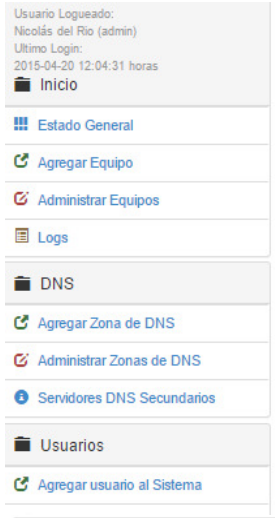
```
/usr/bin/nsupdate -D -y
tm:xq6tQuPF8b8NAVlUYtko4xclMN57eeJXHiiMnY4y67+NjUv0iXnGST0QId
CSlIlN9ClzPEKpbDHsoUkWPnMhJw==
server localhost
zone ejemplo.com.ar
prereq yxdomain www.ejemplo.com.ar
update delete www.ejemplo.com.ar A
update add www.ejemplo.com.ar 60 A 190.40.66.10
send
```

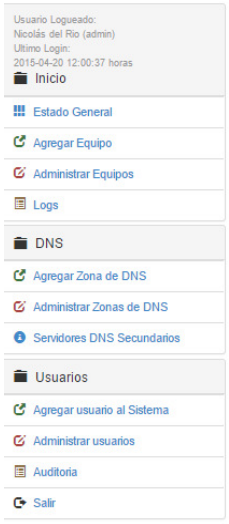
- Toma decisiones y permite que la herramienta escale proveyendo funcionalidad adicional con simples modificaciones

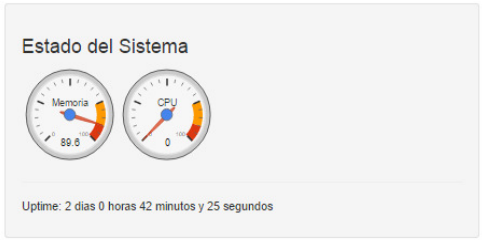
La elección del Valor TTL y tiempos de Expiración de Zonas Un Factor Clave…

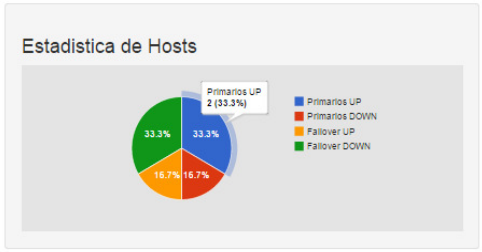
- El valor TTL en las respuestas del DNS debe ser bajo de modo tal de que las consultas no sean cacheadas por largos lapsos de tiempo → Mayor carga en DNS
- El valor del campo expire en la cabecera de las zonas de DNS debe ser alto, de modo que el secundario no expire las zonas en caso de no poder comunicarse con el primario
- Encontrar un equilibrio es fundamental
- Los sistemas operativos no exportan a las aplicaciones el valor de TTL devuelto, con lo cual utilizan valores aleatorios:
  - InterNet Explorer / Edge: 30 minutos / 60 segundos
  - Google Chrome: 60 segundos
  - Mozilla Firefox: 60 segundos

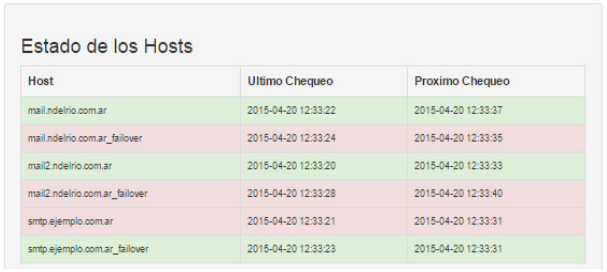
Propuesta de Administración de Tráfico de Red
La Aplicación











Una Pequeña demostración

Conclusiones Finales

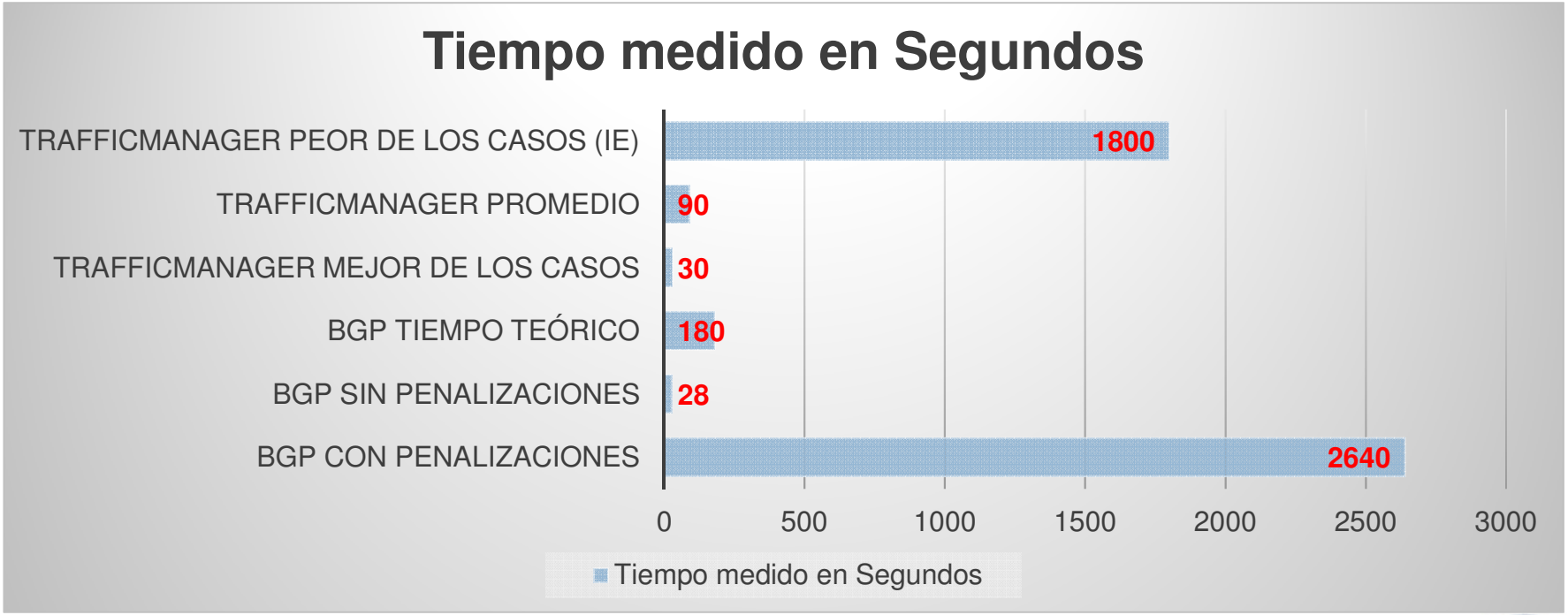
- La alta disponibilidad en el acceso, es un factor indispensable en cualquier infraestructura de servicios
- El protocolo BGP permite la conmutación del tráfico de red hacia enlaces disponibles se manera natural ante la detección de caídas
  - Los tiempos de convergencia son razonables
  - La aplicación de penalizaciones puede generar resultados indeseables
  - Los requisitos para su implementación pueden causar que la solución sea poco accesible a usuarios finales

Conclusiones Finales Cont.

- TrafficManager provee una aproximación a la alta disponibilidad en el acceso a través de la utilización de DNS y chequeos de disponibilidad
  - No puede ser considerado un reemplazo a BGP, pero sí una alternativa
  - Los requisitos para su implementación son mínimos. Lo que hace accesible la solución
  - Los tiempos de convergencia son similares a los de BGP e inclusive mejores si se contempla el uso de penalizaciones
  - Permite realizar conmutaciones por bloques inferiores a los de BGP
  - Su utilización en conjunto con BGP provee un gran potencial

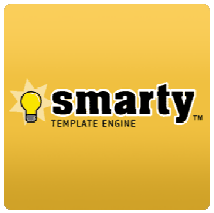
CONCLUSIONES FINALES CONT.

- Tiempos Finales:

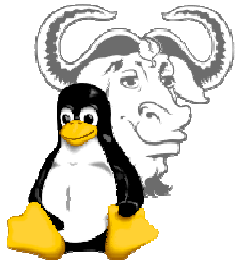


EL TRABAJO EN NÚMEROS

Más de 3000 líneas de código:



Utilización de diversas tecnologías:



Trabajo a Futuro

- Posibilidad de implementar esquemas del tipo Round Robin
- Implementar un proxy de DNS que intercepte las consultas y modifique las respuestas basándose en diverso factores:
  - Destinos de Red
  - Saturación de enlaces, latencia
  - Balanceo
- Implementación de notificaciones
- Adaptar la solución para operar en redes IPv6

Muchas gracias por su tiempo!

¿Preguntas?