Argentina.com : Un estudio de caso

Carlos Horowicz <ch@argentina.com >
Revisión: 43184

FreeBSD is a registered trademark of the FreeBSD Foundation.

CVSup is a registered trademark of John D. Polstra.

Intel, Celeron, EtherExpress, i386, i486, Itanium, Pentium, and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

XFree86 is a trademark of The XFree86 Project, Inc.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this document, and the FreeBSD Project was aware of the trademark claim, the designations have been followed by the "™" or the "®" symbol.

2013-11-13 por hrs.

Tabla de contenidos

1. Introducción 1
2. El desafío 2
3. La solución FreeBSD 3
4. Resultados 6

1. Introducción

Argentina.Com es un ISP argentino con una pequeña infraestructura de menos de 15 empleados y cuya fuente principal de ingresos proviene del negocio del acceso telefónico a redes gratuito. Comenzó a operar en el año 2000 con sólo un servidor para correo y chat.

Desde entonces ha crecido su presencia en un mercado argentino de acceso telefónico a redes que genera unos 45.000 millones de minutos anualmente. Su producto más famoso proporciona a cerca de medio millón de usuarios correo gratuito con webmail, POP3 y

acceso SMPT, junto con 300M de espacio de disco. Hacia el final de 2002 había alrededor de 50.000 usuarios de correo. Después de dos años y medio de reingeniería y de sólidas mejoras técnicas este ISP ha crecido en un factor de 3 en términos de facturación, y en un factor de 10 en cuanto a la base de usuarios de correo.

Nuestros competidores en el mercado argentino de acceso telefónico incluyen a Fullzero (filial perteneciente a Clarin Media Group), Alternativa Gratis y Tutopia, este último fundado por IFX y promocionado por Hotmail. Algunos de estos grandes competidores comenzaron sus respectivos negocios de acceso telefónico con inversiones multimillonarias y con campañas de publicidad agresivas en televisión e Internet. Argentina.Com no utiliza este tipo de publicidad. Ha alcanzado la cuarta posición con un 8% de cuota de mercado durante los dos últimos años gracias a un calidad de servicio superior.

En Argentina y en Latinoamérica en general las personas que no poseen ordenador personal van a los llamados "locutorios" (centros de Internet), donde por unos pocos pesos pueden utilizar un ordenador conectado a Internet y donde normalmente leen y escriben correos electrónicos a través de portales populares como Hotmail, Yahoo! o Argentina.Com.

Debido a los limitados recursos financieros disponibles, Argentina.Com decidió invertir en un nuevo sistema de correo en vez de darse publicidad a través de los medios. Esta decisión estratégica abre las puertas a un futuro negocio en el campo del correo corporativo y de pago.

2. El desafío

El desafío principal para Argentina.Com es alcanzar un tiempo de vida para el servicio de acceso telefónico a redes de al menos 99.95%, o menos de 5 horas de caídas al año. Debido a la alta rotación y volatilidad que existe en este negocio, las cosas deben funcionar correctamente para que el usuario no cambie -voluntariamente o no- de proveedor de acceso a internet o de número de teléfono utilizado para conectarse. El negocio del "dialup", como se le conoce en su denominación inglesa, requiere una estructura de soporte para tratar con las grandes operadoras de telecomunicaciones problemas telefónicos y de calidad de servicio, junto con una infraestructura técnica donde la latencia y la pérdida de paquetes deben minimizarse debido a la naturaleza UDP de los servicio de Radius y DNS, y donde el DNS recursivo debería estar siempre disponible.

Esto tambíen implica tener un tiempo de vida alto en los servicio de POP3 y SMTP, junto con el servicio de webmail. Para POP3 y SMTP se estimó la necesidad de "uptime" igual que para el servicio de "dialup", mientras que para el servicio de webmail se pensó en un porcentaje de 99.5%, lo que significa alrededor de dos dias por año sin servicio o de caída.

Decidimos migrar el correo a una solución propietaria de código abierto que debería ser horizontalmente escalable y cuyos sistemas antivirus y antispam pudieran soportar más de un único tipo de "backend" o de almacenamiento de correos.

La feroz competencia en el mercado del correo electrónico gratuito, principalmente iniciada por las recientes mejoras introducidas por Hotmail, Yahoo! y Gmail, hacían necesario diseñar el nuevo sistema con al menos 300M de espacio de usuario en disco para cada usuario, pero a un coste inferior a 3 dólares americanos por GB incluyendo cierto grado de redundancia. Hay que tener en cuenta que el hardware que puede disponerse en "rack" es difícil de encontrar en Argentina y que resulta ser entre un 30 y un 40% más caro que en los EEUU. Nuestro líquido financiero para adquisición de equipos en dos años fue de 75.000 dólares americanos, lo cual es una fracción muy pequeña de las inversiones acometidas por nuestros competidores directos.

Respecto al servicio antispam, era necesario desarrollar un producto que pudiera competir con los sistemas ofrecidos por los grandes. Dadas las hostiles condiciones que impone la existencia del spam (ataques de diccionario, spams con alto grado de ofuscación y refinamiento, "phishing", troyanos, correos-bomba, etc.) resulta muy complicado alcanzar tiempos de "uptime" excelentes y al mismo tiempo repeler dichos ataques. Uno debe también ser cuidadoso para que el usuario no pierda correos debido a falsos positivos en la estrategia de clasificación, para que no se le inunde con spam o notificaciones de spam y para que el correo peligroso no alcance la carpeta de entrada de los usuarios. Por último, el sistema de correo debe protegerse para que los "spammers" no lo utilicen en su provecho para enviar spam.

El paradigma del código abierto normalmente requiere la adquisición de grandes equipos de administradores de sistemas, operadores y programadores que se encarguen de aplicar parches, corregir "bugs" e integrar plataformas. El paradigma opuesto es también costoso debido a las caras licencias de software, la necesidad de hardware cada vez más caro y debido al elevado número de empleados encargados de proporcionar soporte. Así que el desafío era encontrar la mezcla correcta entre recursos monetarios y humanos escasos, alta estabilidad y grado de predicción, y un desarrollo rápido y fiable. En Buenos Aires resulta difícil encontrar profesionales de las ciencias de la computación bien entrenados, la mayoría de los cuales viven y trabajan en el extranjero, mientras que los restantes poseen trabajos estables dentro de las instituciones del gobierno o en grades compañías.

3. La solución FreeBSD

3.1. Introducción

A comienzos de 2003 teníamos un sistema de correo CriticalPath bajo Solaris x86 y una máquina Redhat para SMTP, Radius y DNS. Los servicios de DNS y Radius se caían constantemente y estábamos luchando con colas enormes de correo electrónico. Hubo un intento de instalar CriticalPath para Linux en Redhat en una máquina Intel con una tarjeta Megaraid, pero la latencia del disco era enorme y la aplicación de correo no llegó a funcionar.

El primer paso realizado hacia la "solución FreeBSD" consitió en migrar este hardware y software comercial a FreeBSD 4.8 con la ayuda de la emulación Linux.

3.2. La elección de FreeBSD

El sistema operativo FreeBSD goza de una merecida fama de por su gran estabilidad, junto con su pragmatismo y sentido común a la hora de poner aplicaciones "on-line" gracias a su excelente colección de Ports. Nosotros consideramos su proceso de generación de releases muy sencillo de entender, además de que la comunidad de usuarios de las listas oficiales de correo electrónico mantiene un estilo educado y civilizado cuando ayudan o leen los problemas de otros usuarios y sus soluciones.

Otra característica importante es su rápida implantación. Afortunadamente pudimos establecer nuestra política de instalación de SO alrededor de las capacidades predefinidas de FreeBSD. En una compañía pequeña algunas veces necesitas ir corriendo a un centro de datos y rápidamente levantar un servidor para proporcionar algún servicio. En los dos últimos años, Argentina.Com adquirió alrededor de cuarenta servidores, la mayoría Pentium IV pero también varios Xeon duales y unos cuantos Opteron duales para ubicarlos en los centros de datos donde tenemos los contratos de operaciones de "hosting" y de acceso telefónico a redes. Todos ellos ejecutan FreeBSD, desde 4.8 (un par de ellos con dos años de "uptime" y cero problemas) hasta 6.0-BETA2.

La política general que tenemos para con el sistema operativo consiste en intentar llevar a todos los servidores a la rama de código estable de una forma periódica utilizando `RELENG_4`, `RELENG_5` y ahora `RELENG_6`. Estas operaciones nos permiten estar más preparados ante posibles amenazas de seguridad a nivel del sistema operativo o del software base del mismo, especialmente en los servidores web.

3.3. Reingeniería básica

El prime paso de reingeniería fue poner en funcionamiento dos máquinas FreeBSD 4.8 cuya única tarea iba a consistir en ser DNS autorizados para todos nuestros dominios. El software elegido resultó ser BIND9. Estas máquinas se ubicaron en diferentes centros de datos, cuidándonos de asegurar una buena latencia entre ellos para evitar problemas en transferencias de zonas, haciendo posible tratar con TTLs entre 60 y 600 segundos para así poseer unos mejores márgenes de reacción en caso de problemas.

El segundo paso consistió en desplegar dos máquinas más del mismo tipo, también en distintos centros de datos, para sólo servir Radius y DNS recursivo. Los servidores de acceso de red ("Network Access Servers" o NAS) de los operadores de telecomunicaciones se configuraron para enviar las peticiones de autorizació y "accounting" de Radius hacia los nuestros, y también para asignar dichos DNS recursivos a nuestros usuarios de acceso telefónico.

La tercera "regla de oro" consiste en no poner jamás en la misma máquina el servicio de entrada y salida de correo SMTP. Desplegamos máquinas FreeBSD distintas utilizando Postfix tanto para la entrada como para la salida.

3.4. Migración del correo

La migración del correo requería una planificación cuidadosa debido al hecho de que íbamos a a migrar tanto los frontales como los "backends". El primer paso fue construir un sistema perimetral antispam y antivirus con FreeBSD 4.x y 5.x basado en postfix, amavisd-new, clamav y SpamAssassin. Estos sistemas iban a entregar correos tanto a los antiguos sistemas como a los nuevos hasta que el nuevo "backend" estuviera en funcionamiento. Entre tanto añadimos pequeñas máquinas FreeBSD para incrementar el "spool" de correo de CriticalPath sin ningún problema.

En la primera línea de la entrada de correo pusimos varios MX del dominio Argentina.com para filtrar ataques de diccionario (intentos de reenviar correo a usuarios no existentes) además de una *lista negra* derivada de SURBL que resultó no dar casi ningún falso positivo. Los correos eran multiplexados hacia un cluster de Xeon duales y Opteron duales donde ejecutábamos amavisd-new junto con un filtrado basado en listas blancas y negras basado en MySQL. Descartamos el uso de Bayes y Autowhitelisting en un nivel global debido a las grandes cantidades de falsos positivos y de falsos negativos que proporcionaban. En su lugar definimos unos cuantos niveles de spam de menos a más tolerante, cada uno con niveles de corte y de descarte. A cada correo electrónico recibido el sistema le asigna una determinada puntuación. Los correos con una puntuación por debajo de la puntuación asociada con el nivel de corte establecido por el nivel de spam pueden continuar hasta la bandeja de entrada del usuario. Los correos entre el nivel de corte y el nivel de descarte se envían a una carpeta del usuario denominada Spam, y por último aquellos correos por encima del nivel de corte se descartan porque se considera una situación muy evidente de spam. En aras de la simplicidad, se asocian de forma transparente los correos almacenados en la libreta de direcciones con el sistema antispam, colocándolos en las "listas blancas" de forma automática.

Con la introducción de Spamassassin 3.x, el tráfico de DNS utilizado para preguntar a las listas negras de correo creció considerablemente, de tal forma que firmamos acuerdos con SpamCop, Spamhaus y SURBL para instalar réplicas públicas de sus bases de datos en nuestro equipo FreeBSD. Gracias a estas réplicas, que nos costaron entre 1 y 2Mbps de tráfico, fuimos capaces de reducir drásticamente la latencia de Spamassassin.

En un tercer nivel nos encontramos con la entrega a los buzones de correo de los usuarios. Tan pronto como nos pusimos a contruir el nuevo "backend" Cyrus-Imap con autentificación MySQL, nos dimos cuenta de que necesitábamos multiplexar el correo de entrada a los usuarios en los formatos de los buzones nuevos y antiguos. Finalmente, conseguimos migrar cientos de miles de correos hacia la nueva arquitectura Cyrus utilizando una fenomenal herramienta denominada imapsync, que se puede instalar directamente desde los ports. También instalamos perdition (un proxy de POP3 y de IMAP) entre medias para asegurar una migraci´n transparente y permitir la distribución entre distintos servidores de los buzones. En resumen, toda la información de localización de un buzón de usuario está en MySQL, y dicha información se encuentra disponible para todo el software que forma parte de la cadena.

Respecto al hardware para el espacio de disco, actualmente utilizamos siete máquinas FreeBSD con Cyrus-Imap de distinto hardware. El mayor es un Pentium IV con 4G de RAM y tarjetas 3ware con 12 bahías extraíbles en caliente, organizadas en 3 unidades RAID-5 de 1 Terabyte cada una. El software 3ware envía un correo electrónico al administrador cuando el RAID se degrada (en la mayoría de los casos se trata de errores de disco) y es posible reconstruir el RAID con el sistema a pleno rendimiento. Utilizamos smartmontools en los casos en los que hay menor redundancia, para disponer inmediatamente de alertas de discos con problemas de temperatura o de fallos de "selftests".

Como software de correo web elegimos un producto comercial denominado Atmail, disponible con sus fuentes en perl y que utiliza mod_perl. Bajo FreeBSD resulta extremadamente sencillo gestionar los módulos de perl y no es necesario usar la "shell" de CPAN; únicamente hay que seleccionar el port correcto y ejecutar "make install". Tras varios meses de trabajo de integración pudimos integrar la parte cliente de Atmail que habla IMAP con nuestros "backends". Tuvimos que modificar algunas partes del código para adaptar el producto a nuestro entorno libre y para hacerlo compatible con nuestro perímetro antispam antivirus, además de aplicar nuestras personalizaciones y traducciones.

3.5. Migración web

Con la adopción de FreeBSD no hubo que realizar ningún esfuerzo adicional para tener en ejecució en cuestion de minutos los entornos de Apache, PHP y MySQL. Incluso las actualizaciones de PHP4 a PHP5 se efectuaron sin problemas. El sistema de ports nos resultó una vez más extremadamente útil y nos permitió hacer cosas como comprimir los contenidos de texto y de html de Apache utilizando unas pocas líneas de documentación. Además, hemos experimentado un rendimiento excelente y una estabilidad y "uptimes" extraordinarios.

4. Resultados

Conseguimos implantar una arquitectura de correo electrónico basada en FreeBSD que es escalable horizontalmente, utilizando 3 Terabytes de almacenamiento basado en servidores Intel incurriendo en un coste de 3 dólares por Gigabyte con redundancia.

La gran estabilidad alcanzada permitió a Argentina.com explorar otros campos como el "hosting" para terceros y el "housing" con presencia en los centros de datos argentinos.

Ahora ofrecemos también acceso telefónico a redes corporativas para usuarios de "roaming" y Perú gracias a nuestra presencia y a los acuerdos suscritos con la mayoría de los operadores de telecomunicaciones. Entre nuestros clientes indirectos se encuentran las principales compañías americanas como Ford, Exxon y Reuters. También estamos en el negocio del acceso telefónico a redes en Brasil, Chile, Colombia y Panamá.