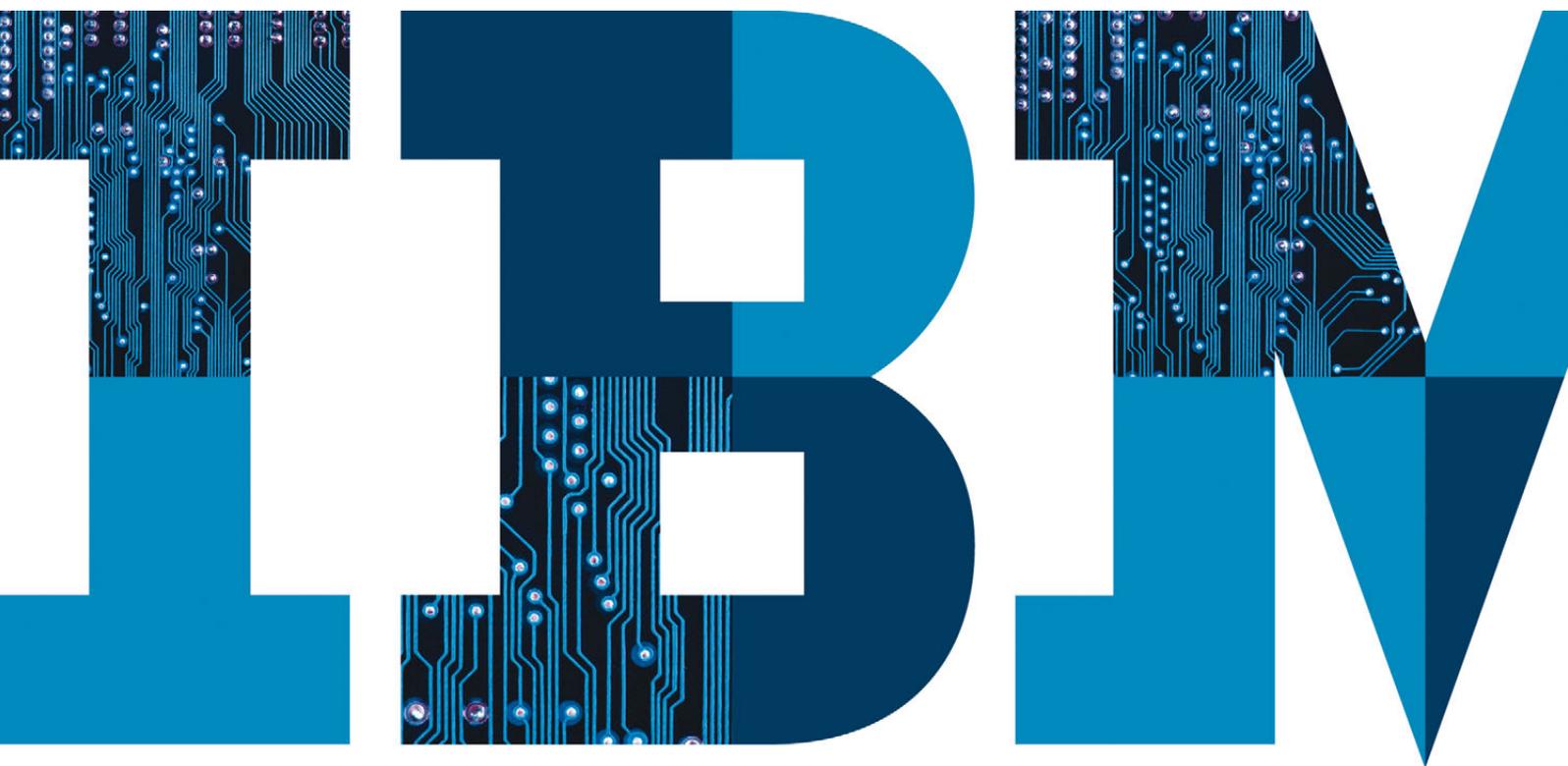


Watson: un sistema diseñado para obtener respuestas

El futuro del diseño de sistemas optimizado de cargas de trabajo



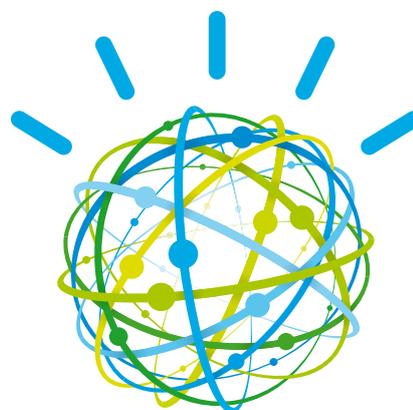
Resumen ejecutivo

Durante el último siglo, IBM ha logrado numerosos avances científicos gracias a su compromiso por la investigación y su tradición de grandes retos. Estos grandes retos (como el Deep Blue, diseñado para hacer frente al campeón mundial de ajedrez Gary Kasparov), han conseguido dar un impulso a la ciencia como nunca antes. Watson es el último gran reto en investigación de IBM; concebido para fomentar la ciencia del procesamiento de lenguajes naturales a través de avances en tecnología de preguntas y respuestas.

Watson es un sistema optimizado de cargas de trabajo fundamentado en la arquitectura IBM DeepQA ejecutada en un clúster de servidores basados en procesadores IBM POWER7. Después de cuatro años de intensas investigaciones y progresos por parte del equipo de IBM, Watson participó en el programa televisivo Jeopardy! en febrero de 2011, compitiendo al mismo nivel que expertos humanos en términos de precisión, confianza y rapidez frente dos de los más conocidos y exitosos ganadores del programa, Ken Jennings y Brad Rutter. Este libro blanco explica el diseño del sistema optimizado de cargas de trabajo de Watson, cómo se ha convertido en el emblema del futuro del diseño de sistemas, y por qué representa un nuevo paradigma informático.

Jeopardy!, el reto de IBM

En 1997, Deep Blue, el sistema informático desarrollado por IBM Research para jugar al ajedrez, atrajo la atención de todo el mundo al competir de manera exitosa contra el campeón mundial de ajedrez Gary Kasparov. Se trataba de la culminación de un gran reto que sirvió para impulsar a la ciencia de la computación de una forma que creó un sobresaliente interés popular.



Hoy en día, con las empresas capturando cada vez más información empresarial crítica en documentación de lenguaje natural, existe un interés creciente por los sistemas optimizados de cargas de trabajo que analizan en profundidad el contenido de preguntas de lenguaje natural para responder a esas preguntas con precisión. Los avances en tecnología de búsqueda de respuestas (QA, por sus siglas en inglés) ayudarán cada vez más a los profesionales de los servicios de asistencia a la hora de tomar decisiones críticas y oportunas en áreas como la sanidad, la inteligencia empresarial (BI, por sus siglas en inglés), el descubrimiento de conocimientos, la gestión del conocimiento empresarial y la atención al cliente.

Teniendo en cuenta la QA, IBM se propuso el reto de crear un sistema informático llamado 'Watson' (en honor de Thomas J. Watson, fundador de IBM), cuyo fin era competir en tiempo real y al mismo nivel que campeones humanos en el concurso de la televisión americana *Jeopardy!*. El programa, que se ha emitido en los EE. UU. durante más de 25 años, enfrenta a tres concursantes humanos entre sí con el objetivo de responder preguntas ricas en lenguaje natural acerca de una amplia variedad de materias, con penalizaciones por cada respuesta incorrecta.



En esta competición entre tres personas, la confianza, la precisión y la rapidez de respuesta son de vital importancia, ya que los concursantes deben hallar la respuesta durante los pocos segundos que tarda el presentador en leer la pista. Para competir en este juego al mismo nivel que los campeones humanos, un sistema informático necesitaría responder alrededor del 70% de las preguntas realizadas, con una precisión superior al 80% de tres segundos o menos.

Watson representa un gran avance en el diseño y análisis de sistemas. Ejecuta tecnología DeepQA de IBM, un nuevo tipo de capacidad de análisis que permite realizar miles de tareas simultáneas en segundos para proporcionar respuestas precisas a las preguntas. Watson, que emplea la tecnología de procesadores IBM POWER7, es un ejemplo de las complejas cargas de trabajo de análisis que son cada vez más comunes y críticas para la competitividad y el éxito empresarial en el entorno de uso intensivo de datos de hoy en día.

Watson compitió contra dos de los concursantes más conocidos y exitosos de *Jeopardy!*, Ken Jennings y Brad Rutter, en un concurso al mejor de dos programas emitidos durante tres noches consecutivas a partir del 14 de febrero de 2011.

IBM DeepQA

DeepQA es una arquitectura probabilística masivamente paralela basada en evidencias. Para el reto de *Jeopardy!* se utilizaron más de 100 técnicas diferentes para analizar el lenguaje natural, identificar fuentes, buscar y generar hipótesis, hallar y conseguir evidencias, y fusionar y clasificar hipótesis. Mucho más importante que cualquier técnica es la manera en la que estas técnicas se combinan en DeepQA, de tal forma que los enfoques solapados pueden mostrar sus puntos fuertes para brindar y contribuir con las mejoras en precisión, confianza o rapidez.



DeepQA es una arquitectura con una metodología de acompañamiento, si bien no está específicamente pensada para el reto de *Jeopardy!*. IBM ha comenzado a adaptarla a diferentes aplicaciones empresariales y a problemas relativos a retos exploratorios adicionales que incluyen medicina, búsqueda empresarial y juegos.

Los principios generales en DeepQA son:

1. **Paralelismo total:** explora el paralelismo total teniendo en cuenta múltiples interpretaciones e hipótesis.
2. **Numerosos expertos:** facilita la integración, aplicación y evaluación contextual de una amplia gama de análisis de contenido y preguntas probabilísticas de acoplamiento flexible.
3. **Estimación de confianza dominante:** ningún componente individual se compromete a una respuesta; todos los componentes producen funciones y confianzas asociadas, logrando diferentes interpretaciones del contenido y las preguntas. Un sustrato de procesamiento de confianzas subyacentes muestra cómo apilar y combinar las puntuaciones.
4. **Integración de conocimiento profundo y superficial:** equilibra el uso de semántica estricta y semántica superficial, aprovechando numerosas ontologías formadas de forma flexible.

Rapidez y ampliación

DeepQA se ha desarrollado utilizando Apache UIMA (Arquitectura para la administración de información no estructurada), una implementación estructural de la UIMA. La UIMA se diseñó para admitir la interoperabilidad y la ampliación de aplicaciones de análisis de multimodalidad y texto. Todos los componentes en DeepQA están aplicados como *anotadores de UIMA*. Se trata de componentes que analizan texto y producen *anotaciones* o afirmaciones sobre el texto. Watson ha evolucionado con el tiempo, contando hoy en día el sistema con cientos de componentes. UIMA facilitó una rápida integración, prueba y evaluación de los componentes.

Las primeras implementaciones de Watson se ejecutaban en un procesador individual, el cual requería dos horas para responder a una única pregunta. La computación de la DeepQA es embarzosamente paralela; sin embargo, se puede dividir en un número de partes independientes, cada una de las cuales se puede ejecutar por un procesador separado. UIMA-AS, parte de Apache UIMA, permite ampliar las aplicaciones de UIMA utilizando mensajes asíncronos. Watson utiliza UIMA-AS

para ampliar a 2.880 cores POWER7 en un clúster de 90 servidores IBM Power 750. UIMA-AS gestiona toda la comunicación entre procesos utilizando el estándar abierto JMS. La implantación de UIMA-AS en POWER7 permitió a Watson proporcionar respuestas en un lapso de uno a seis segundos.

Watson cuenta con unos 200 millones de páginas de contenido de lenguaje natural (equivalente a leer 1 millón de libros). Watson utiliza la estructura Apache Hadoop para facilitar el procesamiento de grandes volúmenes de datos con el fin de crear conjuntos de datos en memoria utilizados durante el tiempo de ejecución. Los anotadores DeepQA UIMA de Watson se implantaron en la estructura MapReduce de Hadoop, distribuyéndolos por procesadores en el clúster. Hadoop contribuye a una óptima utilización de la Unidad de procesamiento central (CPU) y proporciona asimismo herramientas prácticas para la implantación, gestión y supervisión de los procesos de análisis de datos.

Utilización de POWER7

Watson aprovecha el rendimiento de procesamiento masivamente paralelo de sus procesadores POWER7 para ejecutar sus miles de tareas DeepQA de forma simultánea en cores de procesadores individuales. Cada uno de los 90 servidores IBM Power 750 en clústers de Watson incluyen 32 cores POWER7 a 3,55 GHz. Los servidores, que utilizan el sistema operativo Linux®, están alojados en 10 bastidores junto con nodos de entrada/salida (E/S) asociados y centros de comunicaciones. El sistema cuenta con un total combinado de 16 TeraBytes (TB) de memoria y puede funcionar a más de 80 TeraFLOPS (billones de operaciones por segundo).

Gracias a su innovador diseño de procesadores de ocho cores, POWER7 se ajusta perfectamente al procesamiento masivamente paralelo de los algoritmos analíticos de Watson. POWER7 también incluye 500 GigaBytes (GB) de banda ancha de comunicaciones en chip, contribuyendo a una eficiencia excepcional tanto del uso de la memoria como del procesador. Y gracias a que cada servidor incluye 32 cores POWER7 de alto rendimiento con memoria de hasta 512 GB, Power 750 es la plataforma ideal para los procesos Java™ de gran consumo de procesador y memoria de Watson.

El diseño de Watson en servidores Power 750 comercialmente disponibles fue una elección deliberada para garantizar una adopción más rápida de sistemas optimizados en sectores como la sanidad y los servicios financieros. El objetivo era una diferencia fundamental entre Watson y Deep Blue, el cual era un superordenador altamente personalizado. Deep Blue estaba basado en una generación anterior de tecnología de procesadores Power, que incluía un sistema RS/6000 SP de 30 nodos, cada uno con un único procesador POWER2 de 120 MHz. Pero además de los procesadores POWER2 habituales, se mejoró el rendimiento de Deep Blue con 480 chips de procesamiento concebidos específicamente para jugar al ajedrez.

Hoy en día, muchas organizaciones implantan el mismo servidor Power 750 utilizado por Watson en sistemas optimizados que proporcionan tanto análisis complejos como procesamiento de transacciones. La Universidad Rice de Houston, Texas, por ejemplo, utiliza sistemas IBM Power 750 para acelerar la comprensión de la base molecular del cáncer a través de la aplicación de tecnologías de análisis del genoma. Los sistemas POWER7 han proporcionado a Rice una mayor flexibilidad y eficiencia, permitiéndoles dedicarse a una mayor variedad de retos de investigación en un único sistema, algo que antes era impensable. GHY International, una empresa de corretaje de aduanas de Canadá, migró a un nuevo Power 750 con Power AIX, Power i y Power Linux para ofrecer mayor respaldo al creciente compromiso en comercio internacional de sus clientes. Gracias a la virtualización de PowerVM, GHY es ahora capaz de implantar nuevas prestaciones en tan solo cinco minutos para cumplir con las cambiantes necesidades de sus clientes.

Un sistema diseñado para respuestas

Después de cuatro años de intensas investigaciones y progresos por parte del equipo de investigadores de IBM, Watson ha demostrado su capacidad para competir en *Jeopardy!* contra



ganadores del concurso, desempeñando un papel al mismo nivel que expertos humanos en términos de precisión, confianza y rapidez. El proyecto ha superado los campos del análisis de datos no estructurado, el procesamiento de lenguaje natural y el diseño de sistemas optimizados de carga de trabajo. Más allá de *Jeopardy!*, la tecnología que se esconde detrás de Watson se puede adaptar para resolver problemas sociales y empresariales (por ejemplo, el diagnóstico de enfermedades, la gestión de la asistencia técnica en línea y el análisis sintáctico de enormes registros de documentos legales), y para lograr progresos en los diferentes sectores.

La capacidad de Watson para comprender el significado y el contexto del lenguaje humano y para procesar información con rapidez con el fin de hallar las respuestas precisas a preguntas complejas, guarda un enorme potencial para transformar la manera en la que los ordenadores ayudan a las personas a realizar tareas en sus trabajos y vidas personales.

Para más información

Si desea obtener más información sobre Watson, POWER7 y los sistemas optimizados de cargas de trabajo, póngase en contacto con el representante de marketing o Business Partner de IBM, o visite los siguientes sitios web:

- ibm.com/systems/power/advantages/watson
- ibm.com/systems/power



IBM España S.A.

Sta. Hortensia 26-28,
28002 Madrid,
España

El sitio web de IBM está disponible en ibm.com/es

IBM, el logotipo de IBM, ibm.com, AIX, DEEP BLUE, POWER, POWER2, POWER7, PowerVM y RS/6000 son marcas comerciales o marcas registradas comerciales de International Business Machines Corporation en Estados Unidos y/o en otros países. Si estos u otros términos de marcas comerciales de IBM muestran un símbolo de marca comercial (® o ™) la primera vez que aparecen, significa que se trata de marcas comerciales registradas en Estados Unidos o marcas comerciales según derecho consuetudinario propiedad de IBM en el momento en que se publicó esta información. Dichas marcas comerciales también pueden ser marcas registradas o utilizadas en base al derecho consuetudinario en otros países.

Puede consultar una lista actualizada de las marcas comerciales de IBM en Internet, bajo el epígrafe 'Copyright and trademark information', en la dirección ibm.com/legal/copytrade.shtml

Linux es una marca registrada de Linus Torvalds en Estados Unidos y en otros países.

Java y todos los logotipos y marcas comerciales basados en Java son marcas comerciales o marcas comerciales registradas de Oracle y/o sus afiliados.

Otros nombres de empresas, productos y servicios pueden ser marcas comerciales o marcas de servicio de terceros.

Las referencias en esta publicación a productos, programas o servicios de IBM no implican que IBM tenga previsto comercializarlos en todos los países en los que IBM opera.

Las referencias a algún producto, programa o servicio IBM no pretenden dar a entender que sólo puedan utilizarse dichos productos, programas o servicios IBM. En su lugar puede utilizarse cualquier programa, producto o servicio funcionalmente equivalente.

Los productos de hardware de IBM se fabrican con piezas nuevas o con piezas nuevas y usadas revisadas. En algunos casos, es posible que el producto de hardware no sea nuevo y se haya instalado anteriormente. En cualquier caso, se aplican las condiciones de garantía de IBM.

Esta publicación sólo tiene carácter de orientación general. La información está sujeta a cambios sin previo aviso. Póngase en contacto con su representante comercial o distribuidor de IBM local para conocer la información más reciente acerca de los productos y servicios de IBM.

IBM no proporciona consejos legales, contables o de auditoría, ni representa o garantiza que sus productos o servicios cumplan la legislación vigente. Los clientes son responsables de garantizar el cumplimiento de las leyes y normativas, incluidas las leyes y normativas nacionales.

Las fotografías pueden mostrar modelos en fase de diseño.

© Copyright IBM Corporation 2011
Reservados todos los derechos.



Reciclar por favor