

# neo



PONTIFICIA  
UNIVERSIDAD  
CATÓLICA  
DEL PERÚ

Año 1 N° 3  
Del 13 al 26 de  
abril del 2009

## Investigación con **Legión**

A fines del 2008, la Dirección de Informática Académica de la Católica desarrolló un sistema de supercómputo que puede realizar mil millones de operaciones por segundo, lo que permite a sus usuarios ampliar sus posibilidades de estudio y explorar nuevas áreas de investigación.



Por  
JUAN CARLOS  
QUINTANA

**i** Alguna vez te has puesto a pensar cómo los bancos determinan a quién le aprueban un préstamo? Pues realizan un análisis de riesgo crediticio, para lo cual emplean un software de regresión binaria cuyo modelo estadístico evalúa variables como edad, carga familiar e ingresos. Sin embargo, la mayoría de los programas que utilizan las diversas instituciones bancarias en nuestro país, según el Dr. Jorge Luis Bazán, profesor asociado del Departamento de Matemáticas de nuestra Universidad, tienen un retraso de entre diez y quince años.

Al constatar esta problemática, el científico decidió crear un modelo nuevo, para lo cual identificó el problema, estudió los programas disponibles en el mercado y propuso, junto con los profesores Heleno Bolfarine y Marcia D' Branco, de la Universidad de San Pablo

(Brasil), el Skew Probit BBB.

“Actualmente, muchos usuarios utilizan programas de estadística tradicionales. No obstante, estos tienen ciertos problemas y por eso estamos proponiendo una nueva alternativa que ofrece aplicaciones concretas. Por ejemplo, si queremos predecir si participantes de un programa de erradicación de hoja de coca dejarán de asistir al mismo, o saber qué factores de riesgo son importantes para pronosticar una enfermedad, uno va al software comercial y encuentra en el mercado dos o tres tipos de sistemas que puede usar. Pero se han desarrollado muchos más, y uno de esos es el Skew Probit BBB”, señala el Dr. Bazán.

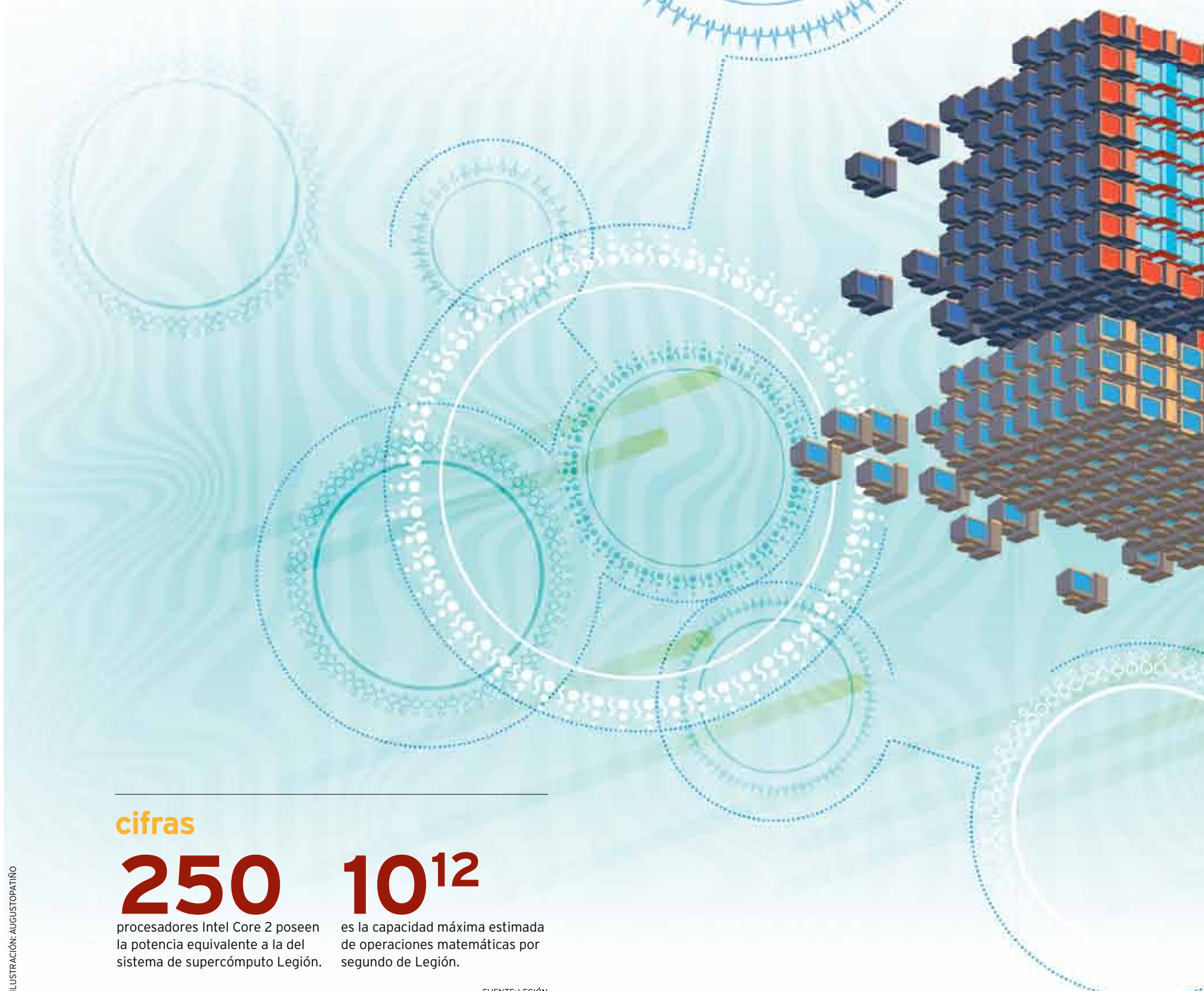
No obstante, para que esta aplicación pudiera ser considerada como una nueva opción, debía de ser validada; es decir, funcionar en diferentes escenarios para comprobar que resultaba mejor que los

La gran ventaja de Legión es que permite al investigador acceder, de manera amigable y desde un navegador web, al BOINC para el envío, la ejecución y el monitoreo de operaciones que requieren grandes cálculos.

modelos tradicionales conseguirlo se tenía que utilizar diversas simulaciones.

“Para formular una propuesta estadística que tiene que demostrarse viable en diferentes escenarios”, resalta el investigador. “Por ejemplo, en los bancos tenemos que probar que funciona para diez personas o para una familia y que aceptará tres o 50. Entonces hay que probar diferentes escenarios que se tienen que representar y esto puede ser posible vía simulación. Así, para demostrar un escenario puedes demostrarlo en unas pocas horas, pero para eso de la programación se tiene que probarlo por lo menos 100 veces. Imagina que demorarías una sola máquina”, dice el Dr. Bazán.

**LEGIÓN AL RESCATE** En esta situación, el Dr. Bazán, en Estadística por



**cifras**

**250**

procesadores Intel Core 2 poseen la potencia equivalente a la del sistema de supercómputo Legión.

**10<sup>12</sup>**

es la capacidad máxima estimada de operaciones matemáticas por segundo de Legión.

ILUSTRACIÓN: AUGUSTOPATINO

FUENTE: LEGIÓN



les. Y para  
an que rea-  
laciones.

una nue-  
dística se  
rar que es  
es escena-  
estigador.

el tema de  
s que com-  
nará para  
a diez mil,  
y variables  
diferentes  
tienen que  
únicamen-  
mulación.

ar un esce-  
orarte dos  
tar seguro  
n debes de  
enos unas  
el tiempo  
utilizando  
”, explica

ATE. An-  
este Doctor  
el Institu-

to de Matemática y Estadística de la Universidad de San Pablo necesitaba una herramienta que le facilitara el trabajo de simulación. Y es que, para realizar una validación en su antigua Pentium IV se demoraba aproximadamente trece horas porque repetía el proceso unas 100 veces, por lo que tenía que dejar un cartelito pegado en el monitor de su computadora, avisando que no la apagaran.

Pero a fines del año pasado la ayuda llegó desde la Dirección de Informática Académica (DIA) de la Católica, que ha desarrollado un sistema de supercómputo al que ha denominado Legión y que posee una capacidad máxima estimada en mil millones de operaciones matemáticas por segundo. Para ello, hace uso del potencial disponible en las computadoras instaladas en todos los laboratorios a cargo de la DIA, lo cual implica que quienes estén haciendo uso

## el dato...

● El equipo de trabajo de la Dirección Informática Académica que desarrolló el sistema de supercómputo Legión estuvo formado por el Ing. Genghis Ríos, coordinador del proyecto, Oscar Díaz Barriga, administrador del servidor BOINC y Martín Iberico Hidalgo, encargado de la programación del sistema Legión.

de dichos equipos no se vean afectados cuando el sistema esté en funcionamiento.

“Actualmente, la DIA tiene a su disposición 450 computadoras que, por lo general, se usan para navegar por Internet y realizar labores de informática. Nosotros hemos ins-



“Este sistema se basa en tecnología de computación en malla denominada *grid computing*, que permite aprovechar los recursos de cómputo y almacenamiento de computadoras dispersas en todo el mundo”.

**ING. GENGHIS RÍOS**  
Coordinador del equipo que desarrolló el sistema Legión para la DIA-PUCP.

talado en cada una de ellas un software libre cuyo nombre es Berkeley Open Infrastructure for Network Computing (BOINC, o la Infraestructura Abierta de Berkeley para la Computación en Red), que opera de manera silenciosa y que no vuelve lentas a las máquinas, pues solo utiliza el 50% de los recursos disponibles en ese momento. Es decir, si un usuario está empleando el 70% de capacidad de una PC, este software utiliza el 50% del 30% disponible; entonces el usuario no lo nota y puede trabajar sin ningún inconveniente”, explica el ingeniero Genghis Ríos, coordinador del equipo que desarrolló este sistema para la DIA.

Sin embargo, para manejar el BOINC se requiere ser experto en sistemas operativos o contar con uno que permita a los investigadores acceder con facilidad. Justamente, la gran ventaja de Legión es que ofrece al investigador

la posibilidad de acceder, de manera amigable y desde un navegador web, al BOINC para el envío, la ejecución y el monitoreo de diversas operaciones que requieren de grandes cálculos.

“Este sistema de supercómputo se basa en tecnología de computación en malla denominada *grid computing*, que permite aprovechar los recursos de cómputo y almacenamiento de computadoras dispersas en todo el mundo. En nuestro caso particular lo que se aprovecha son las 450 computadoras que tenemos en los laboratorios de nuestra Universidad. Por tal motivo, Legión es esa interfaz amigable que oculta al investigador la complejidad del sistema, facilitando el desarrollo de proyectos de investigación que requieran realizar grandes operaciones. Es por eso que lo hemos denominado así, ya que vemos a todas las computadoras como si fue- »

## sabíasque...

- La Infraestructura Abierta de Berkeley para la Computación en Red (BOINC, por sus siglas en inglés) fue desarrollada en 1999 por la Universidad de Berkeley, EEUU, y originalmente estuvo destinada al proyecto SETI@home, un programa que, instalado en las computadoras de cuatro millones de usuarios, analiza señales procedentes del espacio en busca de evidencias de inteligencia extraterrestre; únicamente se activa cuando la máquina no hace nada.
- Boinc está desarrollada bajo la filosofía de ser código abierto y está disponible en varias plataformas incluyendo Windows Vista y diversos tipos de Unix como Linux, FreeBSD o Mac OS X.

## La ciencia necesita tu computadora

Participar activamente en la lucha contra el cáncer, colaborar para detener el calentamiento global o saber si hay vida lejos de nuestra galaxia son posibles -quizá-, sin salir de casa, gastar un sol ni tener conocimientos específicos. ¿Cómo lograrlo? Pues basta con tener una computadora, conexión a Internet e ingresar a la página web de la World Community Grid ([www.worldcommunitygrid.org](http://www.worldcommunitygrid.org)), comunidad sin fines de lucro que utiliza la capacidad ociosa de millones de computadoras de casas, escuelas y oficinas alrededor del mundo para realizar investigaciones en salud y

medio ambiente. “La base de nuestro trabajo es la idea de que la innovación tecnológica combinada con la investigación científica intuitiva y con el voluntariado en gran escala pueden mejorar nuestro mundo. Nuestro éxito depende de individuos que, como usted, contribuyen colectivamente con esta misión, sin fines de lucro, donando su tiempo libre en la PC”, señalan en su página web. Basta con descargar gratuitamente su software. La Universidad de Berkeley, en California, Estados Unidos, tiene una iniciativa similar. Visítala ingresando a la web [boinc.berkeley.edu](http://boinc.berkeley.edu).



» ra, cada una, el soldado de una gran Legión”, resalta el ingeniero Ríos.

No obstante, este sistema no resuelve todos los problemas que necesiten de cómputo intensivo, ya que es necesario trabajar con investigaciones que puedan ser divididas en tareas menores e independientes; es decir, en pequeñas unidades (*workunits*) que puedan ser enviadas a cada una de las computadoras dentro de nuestro campus. Luego de ese procesamiento se tiene un resultado en partes que son sumadas y se obtiene la solución final.

Para utilizar Legión y po-



“Actualmente muchos usuarios utilizan programas de estadística tradicionales. No obstante, estos tienen ciertos problemas y por eso estamos proponiendo una nueva alternativa que ofrece aplicaciones concretas”.

**DR. JORGE LUIS BAZÁN**  
Profesor asociado del Departamento de Matemáticas de nuestra Universidad.

der realizar estos grandes cálculos se recomienda utilizar software libre, dado que el modelo estadístico del profesor Bazán emplea los programas libres R y OpenBUGS, compatibles con el sistema diseñado por la DIA. De esta manera se pudo realizar las simulaciones necesarias para validar, en un menor tiempo, el Skew Probit BBB. Incluso el *paper* al que ha dado lugar este nuevo modelo estadístico ha sido aceptado por una revista especializada para su pronta publicación.

En consecuencia, Legión ofrece a nuestros investigadores la posibilidad de acceder a un sistema de cómputo de alto rendimiento de manera sencilla y económica y se vislumbra como una poderosa herramienta que reducirá los tiempos para validar nuevos modelos estadísticos y permitirá incrementar la producción científica y el posicionamiento académico de nuestra casa de estudios. Además, posibilitará ingresar a nuevos campos de investigación como la predicción climatológica, la química cuántica, la física de altas energías, la exploración petrolera y sísmica, la bioinformática, el análisis de estructuras de proteínas, las simulaciones de fluidos, el manejo intensivo de animaciones e imágenes en 3D, la minería de datos, entre otras áreas.

Sin duda, el inicio de los grandes proyectos de investigación recién empieza. ■

**+información:**

Para conocer más acerca del sistema Legión y cómo utilizarlo, escribe a la Dirección de Informática Académica, al correo electrónico [legion@puccp.edu.pe](mailto:legion@puccp.edu.pe)

**entrevista**

**ING. JOAQUÍN GUERRERO**  
Director de la Dirección de Informática Académica (DIA-puCCp)

# “No hay otra universidad en el país que tenga algo parecido a Legión”

YANINA PATRICIO



**¿Qué significa para la Universidad contar con este sistema de supercómputo?**

Como representante del Perú ante la Cooperación Latinoamericana de Redes Avanzadas (CLARA), he tenido la oportunidad de ver las diferentes investigaciones que se llevan a cabo en Latinoamérica; incluso en países como Ecuador, Colombia y Venezuela hay un desarrollo importante en el ámbito de la investigación. No obstante, es triste ver que en el Perú aún no despegamos en esa área. En estos países ya han empezado a trabajar con lo que es la computación en grandes volúmenes. Actualmente, está lo que se conoce como computación en *grid*, que implica la unión de varias computadoras ubicadas en distintas partes del mundo a fin de sumar las capacidades que tienen. Por ende, como en el mundo se está optando por esta tecnología y por el trabajo colaborativo, no podíamos permanecer relegados. Por eso, el hecho de que tengamos la posibilidad de hacer grandes cálculos debería representar un gran impulso para la investigación.

**¿Entonces ese sería el principal objetivo de Legión?**

Indudablemente, el incentivar la investigación que, además de tener una relación directa con el bienestar social, es uno de los principales objetivos. Ese es nuestro principal propósito con este siste-

ma de supercómputo: fomentar entre nuestros profesores la investigación, ya que con Legión ellos tienen la posibilidad de trabajar en proyectos de mayor envergadura porque el sistema lo facilita. Además, hay que precisar que es un paso previo para que nuestros estudiosos estén en capacidad de interactuar con colegas de otros países en grandes proyectos como, por ejemplo, en el análisis de datos del Gran Colisionador de Hadrones (Large

las pruebas psicológicas. Es decir, nuestro trabajo es ir conociendo los requerimientos de cada unidad. Ahora bien, el sistema Legión nace como una iniciativa que se enmarca dentro de las funciones de la DIA, pero que es una iniciativa particular, como nacieron todos los proyectos dentro de esta unidad, que se caracteriza por tener mucha gente creativa que está buscando qué es lo que se necesita en el ámbito informático. Nuestro

## “Nuestro principal objetivo con este sistema es fomentar la investigación en nuestros profesores”

Hadron Collider o LHC), entre otras áreas investigativas. Después de lograr este objetivo podemos ofrecer servicios a terceros o firmar convenios con otras instituciones. El futuro, sin duda, es prometedor.

**¿Cómo así decidieron crear esta importante herramienta?**

Parte de las funciones de la DIA es el desarrollo e introducción de tecnologías aplicadas a cada una de las especialidades que posee la Universidad Católica. Tenemos que ir buscando cómo apoyar, por ejemplo, a los psicólogos, y por eso estamos desarrollando medios informáticos para la ejecución de

equipo siempre está al tanto de lo que pasa en el mundo. Así nació este sistema y de seguro aparecerán otras aplicaciones en el futuro.

**Tengo entendido que somos la única universidad peruana que posee un sistema como este.**

Tiene usted razón. No hay otras universidades que tengan algo parecido a Legión, somos la única en el país que lo tiene. De esta manera, esperamos que llegue a nuestros investigadores; son ellos los que deberían explotar esta herramienta que está disponible y que tiene un gran potencial.

**enlweb**



● Visita la página web de Legión, ingresando a <http://legion.puccp.edu.pe/>



● Para informarte acerca de los proyectos a los que te puedes unir mira el video “Únete a Boinc”, en <http://www.youtube.com/watch?v=tzv78mdQ3s4>