

Curso: Algoritmos Paralelos y Computación de Altas Prestaciones

BLAS

Domingo Giménez

Departamento de Informática y Sistemas

Universidad de Murcia, Spain

dis.um.es/~domingo





Motivación

- Muchas veces son la parte más costosa de la resolución del problema
- Identificar rutinas básicas, estandarizar e implementarlas eficientemente:
 - Programación más fácil
 - Mantenimiento más fácil
 - Más portabilidad
 - Códigos eficientes



Motivación

- Un gran número de aplicaciones científicas hacen uso del Álgebra Lineal Numérica:
 - Simulación de moléculas (problemas de valores propios)
 - Econometría (mínimos cuadrados)
 - Radiosidad (sistemas de ecuaciones)
 - Búsqueda de información en web (valores propios)
 - Reconocimiento de caras (valores propios)

Antecedentes históricos

(fuente: Enrique Quintana, UJI)

- 1962: Rounding errors, Wilkinson
- 1965: The Algebraic Eigenvalue Problem, Wilkinson. SVD, Golub, Kahan
- 1969: **NAG**. Strassen
- 1972: **EISPACK**. QZ, Moler, Stewart
- 1973: **BLAS** Report
- 1974: Inicio **LINPACK**. Inicio **BLAS-1**

Antecedentes históricos

(II)

- 1975: ACM Trans. on Math. Soft.
- 1976: **EISPACK 2.0**
- 1977: Fortran 77
- 1978: **LINPACK. BLAS1** en ACM TOMS
- 1980: DV, Cupper
- 1982: Inicio **BLAS-2**
- 1983: Matrix computation, Golub, Van Loan

Antecedentes históricos

(III)



- 1984: **EISPACK 3.0.** netlib
- 1986: Inicio **BLAS3**
- 1987: Inicio **LAPACK.** **BLAS2** en ACM TOMS
- 1990: **BLAS3** en ACM TOMS.
- 1992: **LAPACK 3.0**
- 1993: Inicio **ScaLAPCK**

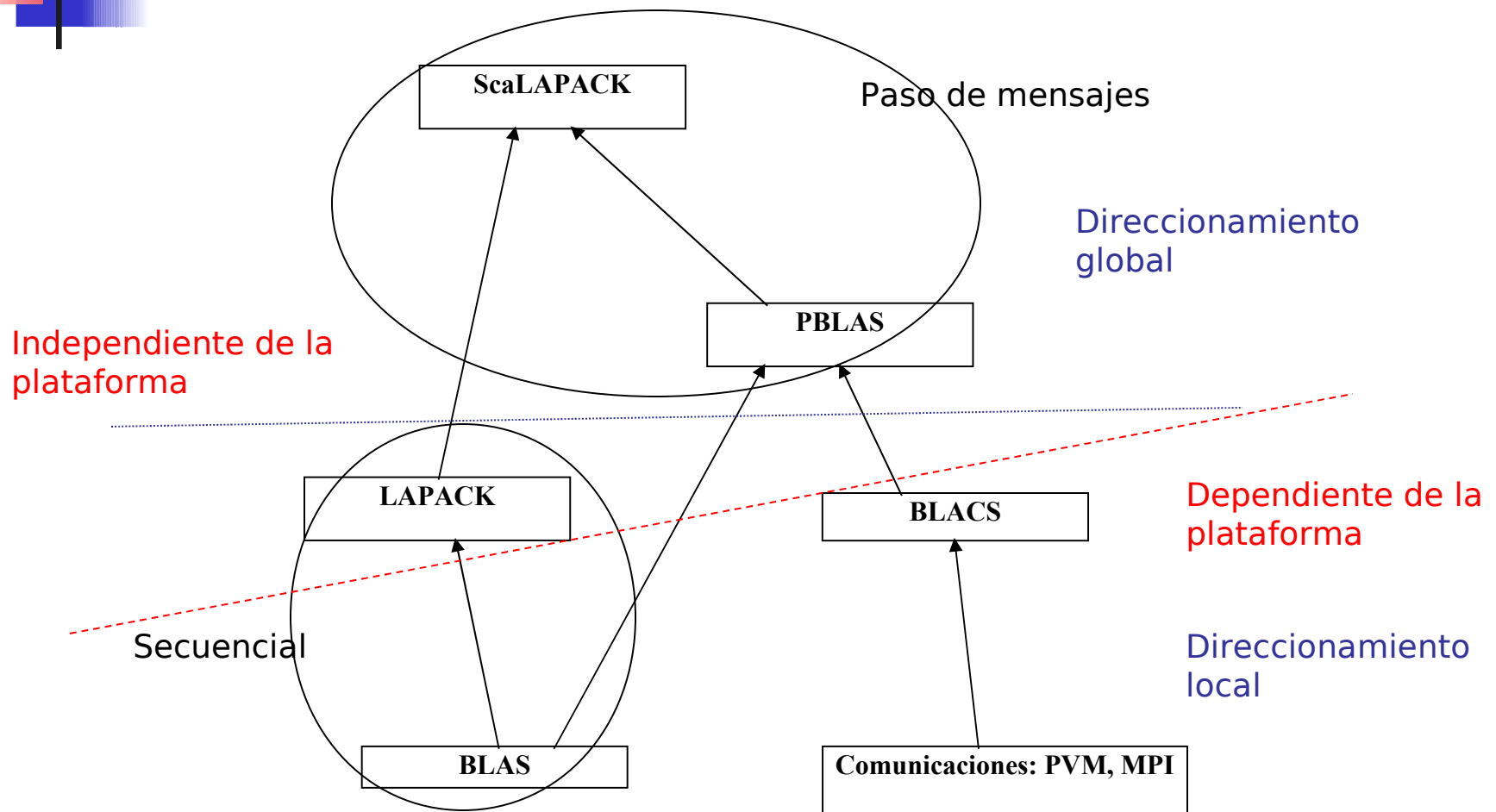
Antecedentes históricos

(IV)



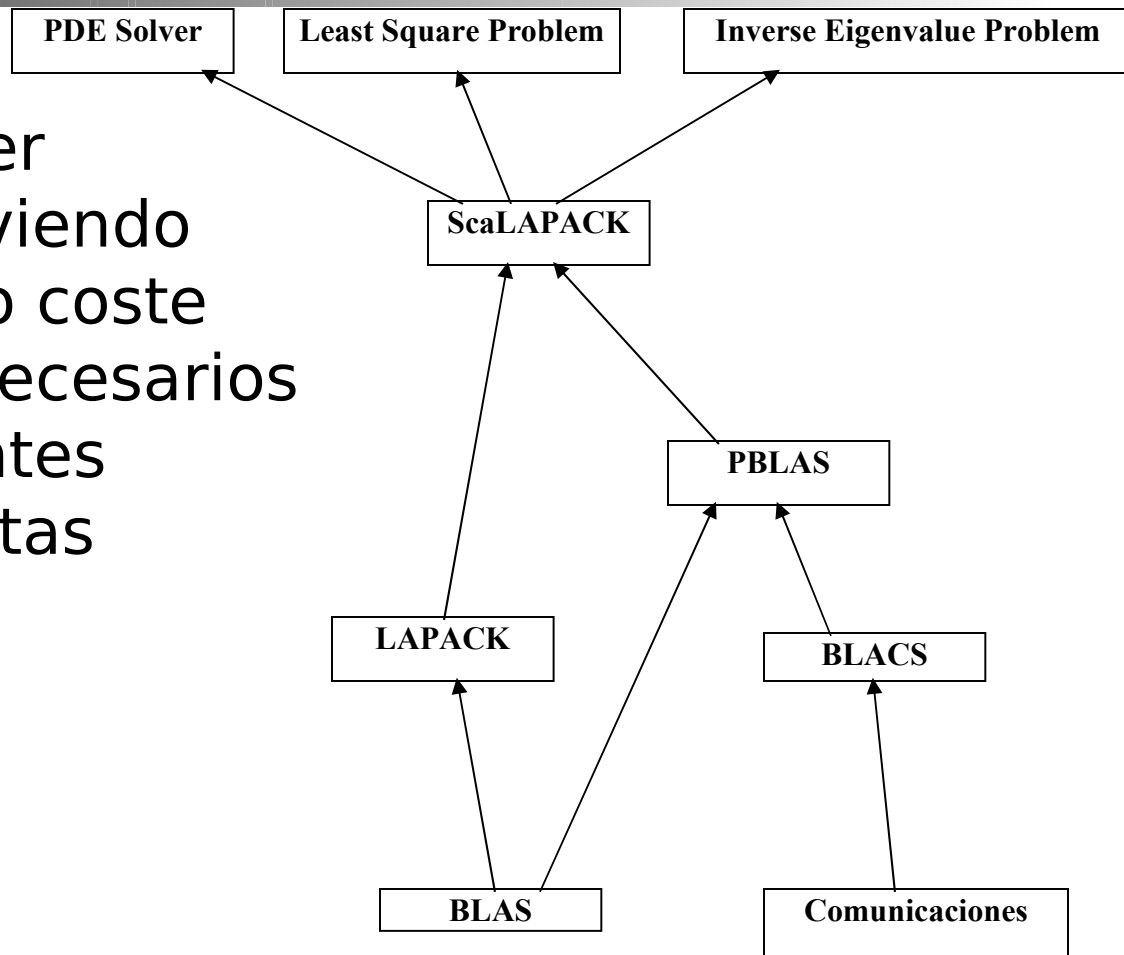
- 1996: ARPACK
- 1997: ScaLAPACK. PLAPACK
- 1999: SuperLU
- 2000: ATLAS. PETSc
- 2002: GotoBLAS
- ...: FLAME. LAPACK07.
HeteroScaLAPACK

Jerarquía de librerías



Jerarquía de librerías

Se puede extender la jerarquía resolviendo problemas de alto coste computacional. Necesarios algoritmos eficientes en sistemas de altas prestaciones.

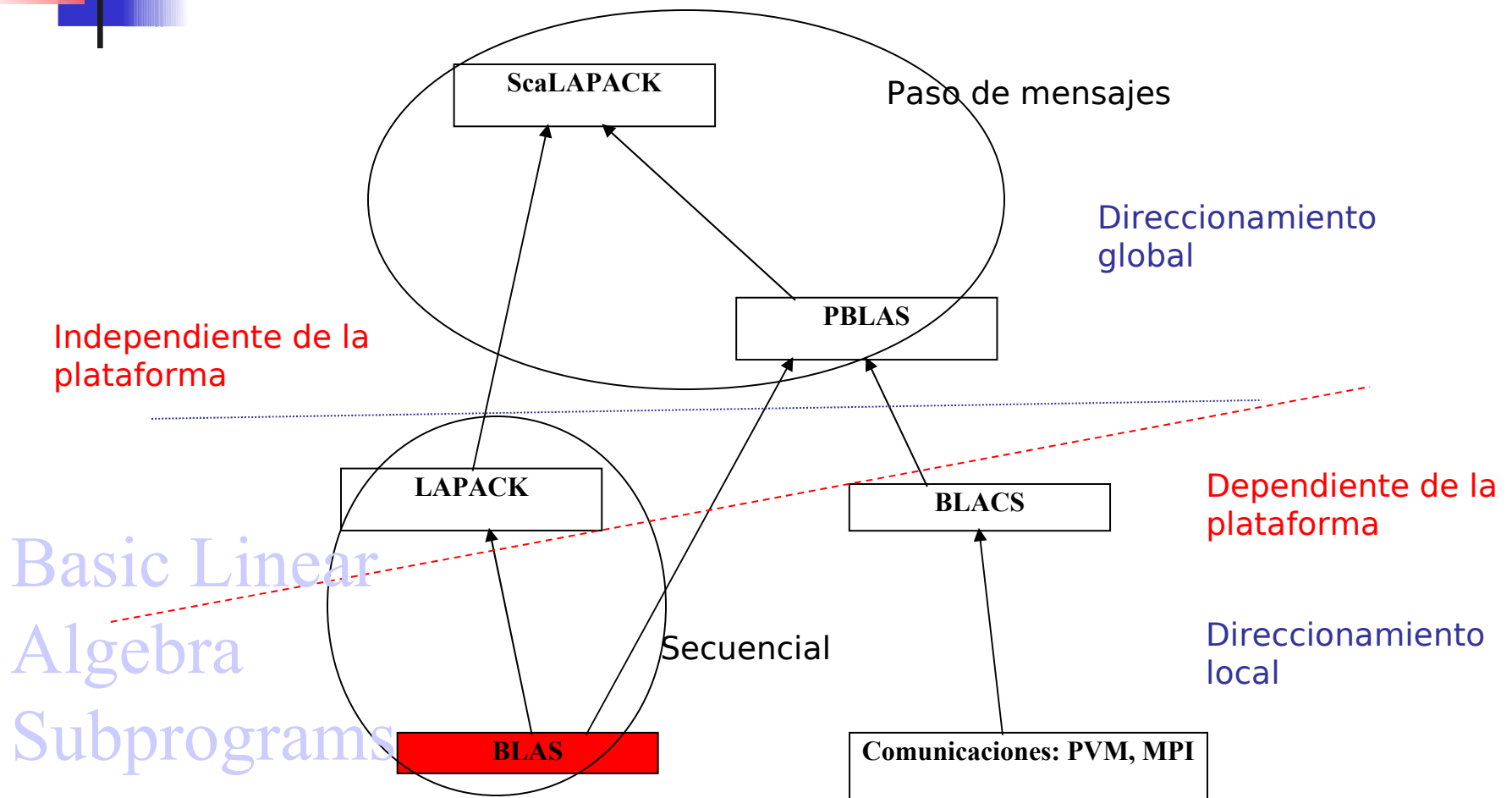




Obteniendo información

- www.netlib.org/liblist.html
- www.netlib.org/utk/people/JackDongarra/la-sw.html

BLAS





BLAS

- Conjunto de rutinas para realizar operaciones básicas sobre vectores y matrices
 - C. L. Lawson, R. J. Hanson, D. Kincaid, and F. T. Krogh, *Basic Linear Algebra Subprograms for FORTRAN usage*, [ACM Trans. Math. Soft., 5 \(1979\)](#), pp. 308--323.
 - J. J. Dongarra, J. Du Croz, S. Hammarling, and R. J. Hanson, *An extended set of FORTRAN Basic Linear Algebra Subprograms*, [ACM Trans. Math. Soft., 14 \(1988\)](#), pp. 1--17.
 - J. J. Dongarra, J. Du Croz, I. S. Duff, and S. Hammarling, *A set of Level 3 Basic Linear Algebra Subprograms*, [ACM Trans. Math. Soft., 16 \(1990\)](#), pp. 1--17.



BLAS

- Hay tres niveles según el coste computacional:

	tipo operaciones	coste computacional	accesos memoria
BLAS1	vector-vector	n	n
BLAS2	matriz-vector	n^2	n^2
BLAS3	matriz-matriz	n^3	n^2



BLAS 1

Level 1 BLAS

	dim	scalar	vector	vector	scalars	5-element array
SUBROUTINE xROTG (A, B, C, S)	
SUBROUTINE xROTMG(D1, D2, A, B,	PARAM)
SUBROUTINE xROT (N,			X, INCX, Y, INCY,		C, S)	
SUBROUTINE xROTM (N,			X, INCX, Y, INCY,			PARAM)
SUBROUTINE xSWAP (N,			X, INCX, Y, INCY)			
SUBROUTINE xSCAL (N,	ALPHA,		X, INCX)			
SUBROUTINE xCOPY (N,			X, INCX, Y, INCY)			
SUBROUTINE xAXPY (N,	ALPHA,		X, INCX, Y, INCY)			
FUNCTION xDOT (N,			X, INCX, Y, INCY)			
FUNCTION xDOTU (N,			X, INCX, Y, INCY)			
FUNCTION xDOTC (N,			X, INCX, Y, INCY)			
FUNCTION xxDOT (N,	ALPHA,		X, INCX, Y, INCY)			
FUNCTION xNRM2 (N,			X, INCX)			
FUNCTION xASUM (N,			X, INCX)			
FUNCTION IxAMAX(N,			X, INCX)			



BLAS 1

Ejemplo `ddot.f`

Calcula el producto escalar de dos vectores

Se puede usar en el bucle más interno de la multiplicación de matrices, dando lugar a una **versión con BLAS 1**

Se compila con (depende del sistema)

`cc -O3 mb1.c -lblas -lm`



BLAS

Formato de las funciones (niveles 2 y 3): XYYZZZ

X: Tipo de datos:

S : REAL

D : DOUBLE PRECISION

C : COMPLEX

Z : DOUBLE COMPLEX

YY: Tipo de matriz: GE, GB, HE, HP, HB, SY, SP, TR, TP, TB

ZZZ: Operación:

MV: productor matriz vector

MM: producto matriz matriz

SV: sistema de ecuaciones ...

BLAS 2

	options	dim	b-width	scalar	matrix	vector	scalar	vector
xGEMV (TRANS,	M, N,		ALPHA,	A, LDA,	X, INCX,	BETA,	Y, INCY)
xHEMV (TRANS,	M, N, KL, KU,		ALPHA,	A, LDA,	X, INCX,	BETA,	Y, INCY)
xHEMV (UPLO,	N,		ALPHA,	A, LDA,	X, INCX,	BETA,	Y, INCY)
xHEMV (UPLO,	N, K,		ALPHA,	A, LDA,	X, INCX,	BETA,	Y, INCY)
xHPMV (UPLO,	N,		ALPHA,	AP,	X, INCX,	BETA,	Y, INCY)
xsymv (UPLO,	N,		ALPHA,	A, LDA,	X, INCX,	BETA,	Y, INCY)
xSEMV (UPLO,	N, K,		ALPHA,	A, LDA,	X, INCX,	BETA,	Y, INCY)
xSPMV (UPLO,	N,		ALPHA,	AP,	X, INCX,	BETA,	Y, INCY)
xTRMV (UPLO, TRANS, DIAG,	N,			A, LDA,	X, INCX)		
xTMV (UPLO, TRANS, DIAG,	N, K,			A, LDA,	X, INCX)		
xTPMV (UPLO, TRANS, DIAG,	N,			AP,	X, INCX)		
xTRSV (UPLO, TRANS, DIAG,	N,			A, LDA,	X, INCX)		
xTSV (UPLO, TRANS, DIAG,	N, K,			A, LDA,	X, INCX)		
xTPSV (UPLO, TRANS, DIAG,	N,			AP,	X, INCX)		

BLAS 2

	options	dim	scalar	vector	vector	matrix
xHER	(M, N,	ALPHA,	X, INCX,	Y, INCY,	A, LDA)
xHERU	(M, N,	ALPHA,	X, INCX,	Y, INCY,	A, LDA)
xHERC	(M, N,	ALPHA,	X, INCX,	Y, INCY,	A, LDA)
xHER	(UPLO,	N,	ALPHA,	X, INCX,		A, LDA)
xHPR	(UPLO,	N,	ALPHA,	X, INCX,		AP)
xHER2	(UPLO,	N,	ALPHA,	X, INCX,	Y, INCY,	A, LDA)
xHPR2	(UPLO,	N,	ALPHA,	X, INCX,	Y, INCY,	AP)
xsyr	(UPLO,	N,	ALPHA,	X, INCX,		A, LDA)
xspr	(UPLO,	N,	ALPHA,	X, INCX,		AP)
xsyr2	(UPLO,	N,	ALPHA,	X, INCX,	Y, INCY,	A, LDA)
xspr2	(UPLO,	N,	ALPHA,	X, INCX,	Y, INCY,	AP)



BLAS 2

Ejemplo `dgemv.f`

Calcula el producto de una matriz por un vector

Se puede usar en el segundo bucle en la multiplicación de matrices, dando lugar a una `versión con BLAS 2`

Se compila con

`cc -O3 mb2.c -lblas -lm`

BLAS 3

options

dim

scalar

matrix

matrix

scalar

matrix

xGEMM (TRANSA, TRANSB, N, M, K, ALPHA, A, LDA, B, LDB, BETA, C, LDC)

xSYMM (SIDE, UPLO, N, M, ALPHA, A, LDA, B, LDB, BETA, C, LDC)

xHEMM (SIDE, UPLO, N, M, ALPHA, A, LDA, B, LDB, BETA, C, LDC)

xSYRK (UPLO, TRANS, N, K, ALPHA, A, LDA, BETA, C, LDC)

xHERK (UPLO, TRANS, N, K, ALPHA, A, LDA, BETA, C, LDC)

xSYRK2K (UPLO, TRANS, N, K, ALPHA, A, LDA, B, LDB, BETA, C, LDC)

xHERK2K (UPLO, TRANS, N, K, ALPHA, A, LDA, B, LDB, BETA, C, LDC)

xTRMM (SIDE, UPLO, TRANSA, DIAG, M, N, ALPHA, A, LDA, B, LDB)

xTRSM (SIDE, UPLO, TRANSA, DIAG, M, N, ALPHA, A, LDA, B, LDB)



BLAS 3

Ejemplo `dgemv.f`

Calcula el producto de una matriz por un vector

Se puede hacer la multiplicación de matrices llamando directamente a la rutina correspondiente de BLAS

Se compila con

```
icc -O3 mb3.c -lgslcblas -lm
```



BLAS

- Multiplicación de matrices (en kefren, pentium 4):

Método\tam	200	400	800
Normal	0.0463	0.7854	7.9686
Blas 1	0.0536	0.8190	8.2311
Blas 2	0.0501	0.5861	5.9997
Blas 3	0.0429	0.6115	4.7252



Versiones de BLAS

- BLAS de referencia: los códigos, los podemos instalar en nuestro sistema, no optimizado
- BLAS propietario: optimizado? por los vendedores para su sistema
 - Intel: mkl
 - IBM: ESSP ...
- GotoBLAS: muy eficiente en algunos casos
- Multitud de versiones libres optimizadas? para distintos sistemas, precompiladas
- ATLAS se autoinstala



Algoritmos por bloques

- En vez de realizar operaciones elemento a elemento realizarlas con bloques de elementos: menos accesos a memoria para el mismo volumen de computación \Rightarrow menor tiempo de ejecución.
- Técnica utilizada desde los años 80. Se utiliza en **LAPACK** para obtener rutinas eficientes independientemente del sistema donde se ejecuten.