

***212***

SNELL & WILMER, L.L.P.  
Alan L. Sullivan (3152)  
Todd M. Shaughnessy (6651)  
Amy F. Sorenson (8947)  
15 West South Temple, Suite 1200  
Salt Lake City, Utah 84101-1004  
Telephone: (801) 257-1900  
Facsimile: (801) 257-1800

CRAVATH, SWAINE & MOORE LLP  
Evan R. Chesler (admitted pro hac vice)  
David R. Marriott (7572)  
Worldwide Plaza  
825 Eighth Avenue  
New York, New York 10019  
Telephone: (212) 474-1000  
Facsimile: (212) 474-3700

*Attorneys for Defendant/Counterclaim-Plaintiff  
International Business Machines Corporation*

FILED IN UNITED STATES DISTRICT  
COURT, DISTRICT OF UTAH

AUG 4 3 2004

BY MARKUS B. ZIMMER, CLERK  
DEPUTY CLERK

IN THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF UTAH

THE SCO GROUP, INC.

Plaintiff/Counterclaim-  
Defendant,

vs.

INTERNATIONAL BUSINESS  
MACHINES CORPORATION,

Defendant/Counterclaim-  
Plaintiff.

**DECLARATION OF  
BRIAN W. KERNIGHAN**

Civil No. 2:03CV0294 DAK

Honorable Dale A. Kimball

Magistrate Judge Brooke Wells

**FILED UNDER SEAL**

**I. Introduction**

1. I have been asked by counsel for IBM to evaluate the opinions set out in the declaration of Sandeep Gupta submitted in opposition to IBM's motion for partial summary judgment of non-infringement with respect to its Linux activities. Specifically, I have been asked to address Mr. Gupta's conclusion that there is "substantial similarity" between certain "routines" and "groupings of code" in Linux and copyrighted works allegedly owned by SCO.
2. My analysis and conclusions are based upon the principles described in *Gates Rubber v. Bando*, 9 F.3d 823 (10<sup>th</sup> Cir. 1993), which I understand to describe the appropriate methodology for determining substantial similarity. My analysis and conclusions are also based upon my experience and expertise in the field of computer science.
3. In summary, I find fundamental errors in Mr. Gupta's conclusions. His conclusions of substantial similarity are flawed because he fails to exclude from comparison unprotectable elements of the allegedly copyrighted code, and he uses an indefensible standard for what qualifies as "substantially similar" code.
4. If unprotectable elements are excluded from the comparison and an appropriate standard of similarity is applied, there is no similarity between the parts of Linux identified by Mr. Gupta and the allegedly copyrighted works.
5. Errors in methodology aside, the code Mr. Gupta identifies as infringing does not constitute a substantial part of the allegedly copyrighted works. SCO purports to hold copyrights in at least several versions of Unix System V source code, which over the years have ranged in size from several hundred thousand lines of code (in early versions)

---

to many millions of lines of code in a current version. The alleged similarities identified by Mr. Gupta amount to less than three hundred lines of code, none of which is qualitatively or quantitatively significant.

6. The next section describes my background and qualifications to address the issues addressed herein. The following section describes the flaws in Mr. Gupta's conclusions and explains why, when the segments of material in Linux to which SCO claims rights are compared to protectable elements of the allegedly copyrighted works, there is no similarity between them.

## **II. Background and Qualifications**

7. I am a professor of Computer Science at Princeton University in Princeton, New Jersey. Exhibit I provides more details of my technical background and experience. I received my undergraduate degree in Engineering Physics from the University of Toronto in 1964, and a PhD in Electrical Engineering from Princeton in 1969.
8. From 1969 to 2000 I was a member of technical staff in the Computing Science Research Center at Bell Labs in Murray Hill, New Jersey, where I worked on programming languages, software tools, programming methodology, and other areas, in the group that developed the Unix operating system, the C and C++ programming languages, and numerous other software systems. My personal research included the development of a number of software systems, some of which are in widespread use today, such as the AWK programming language, the AMPL modeling language, and a variety of document preparation tools. From 1981 to 2000, I was head of the Computing Structures Research department. I was named a Bell Labs Fellow in 1996.

9. I am the co-author of eight books on programming, which have been translated altogether into 24 languages, and over 60 technical articles, and hold four software patents. Together with Dennis Ritchie (one of the two inventors of Unix), I wrote *The C Programming Language* in 1978, which was the first book describing C, and which has since become a standard text on the language. I am also the co-author, with Rob Pike, of *The Unix Programming Environment* and *The Practice of Programming*.
10. I was an advisor to Prentice-Hall for their software series from 1978 to 1990, and have been the advisor for Addison-Wesley's Professional Computing series since 1990. I am an editor for the journal *Software Practice and Engineering*. From 1993 to 1995, I was a member of the computer science grant selection committee for the National Science and Engineering Research Council (Canada). I have served on advisory committees at Stanford University, Carnegie-Mellon University, and Princeton.
11. From 1996 to 1999, I was a member of the National Research Council's Panel for Information Technology, and was co-chair from 1997 to 1999. This panel reviews and assesses the work of the information technology laboratory at the National Institute of Standards and Technology (NIST), in the areas of computing, mathematics, software testing, networking, security, and the like.
12. I am a member of the National Academy of Engineering ("NAE"), to which I was elected in 2002, and I am currently a member of the NAE Peer Committee for the Computer Science and Engineering section.
13. I have not testified as an expert in a trial or deposition in the last four years.

- 
14. I have been retained by counsel for IBM in this lawsuit and am being compensated at a rate of \$400 per hour.

**III. Analysis of the Gupta Declaration**

15. Mr. Gupta concludes that there are six categories of “routines” or “groupings of code” in Linux that are substantially similar to the allegedly copyrighted works: (1) the Read-Copy-Update (“RCU”) routine; (2) the user level synchronization (“ULS”) routines; (3) IPC code; (4) certain “headers and interfaces”; (5) UNIX System V init code; and (6) Executable and Linking Format (“ELF”) code (¶ 3). Mr. Gupta’s conclusions regarding these six categories of routines or groupings of code appear to form the basis of SCO’s contention that IBM’s Linux activities infringe SCO’s alleged copyrights.
16. In *Gates Rubber*, the Tenth Circuit sets out a three-part test, known as the abstraction-filtration-comparison test, for determining whether a computer program (such as Linux) is substantially similar to and therefore may infringe the copyright in another computer program (such as Unix). 9 F.3d at 834-39. While flexible, the test ordinarily contemplates that a court (1) “dissect the program according to its varying levels of generality”, (2) “filter out those elements of the program which are unprotectable” at each level of abstraction; and (3) “compare the remaining protectable elements with the allegedly infringing program” *Id.* at 834.
17. Mr. Gupta does not describe the methodology by which he arrives at his conclusions. Even so, it is clear that Mr. Gupta’s methodology, and therefore his conclusions, are indefensible. Mr. Gupta does not appear to have dissected the allegedly copyrighted

---

works according to their varying levels of generality, but the real problem with his analysis is that he fails to filter unprotectable elements before comparing the works at issue and he employs a mistaken standard of similarity.

18. As is explained in *Gates Rubber*, the unprotectable elements of an allegedly copyrighted work must be filtered out of the analysis before an allegedly infringing work may be found to be substantially similar. 9 F.3d at 838-39. "Filtration should eliminate from the comparison the unprotectable elements of ideas, processes, facts, public domain, information, merger material, scènes à faire material, and other unprotectable elements suggested by the particular facts of the program under examination." *Id.* at 834. To the extent there are similarities between the routines and groupings of code identified by Mr. Gupta, they concern unprotectable elements. They lack originality or constitute ideas, procedures, processes, systems, methods of operation, concepts, principles or discoveries, public domain information, merger material, and/or scènes à faire material.
19. Moreover, substantial similarity may be found, according to the Tenth Circuit, only where "those protectable portions of the original work that have been copied constitute a substantial part of the original work -- i.e., matter that is significant in the plaintiffs' program". *Gates Rubber*, 9 F.3d, at 839. Putting aside Mr. Gupta's failure to filter unprotectable elements, much of what Mr. Gupta is willing to call "substantially similar," or even "identical," plainly is not, even to an untrained eye. In some instances, code that Mr. Gupta discusses as if it were a single block is in fact his own composite of code scattered throughout a large body of code, sometimes even coming from different files.

Finally, the amount of code at issue is an insignificant portion of the whole — not only quantitatively but also qualitatively.

20. Even without employing all potential filters, I have no difficulty concluding that all of the similarities identified by Mr. Gupta either should have been filtered as unprotectable or are simply not similar, much less “identical”.
21. Two allegedly copied regions (described by Mr. Gupta as the “RCU routine” and the “ULS routines”) are unprotectable ideas or processes expressed at a very high level of abstraction, not protectable expression. Mr. Gupta’s own language makes the point, labeling the material at issue a “routine” (¶¶ 3, 5, 10) and “method” (¶¶ 6, 7). He also describes the routines at issue as “perform[ing] the same five acts” (¶ 11) and “includ[ing] the same nine (9) acts” (¶ 36), but each of these acts is generic, far removed from a specific expression. Moreover, the actual code that Mr. Gupta says is “substantially similar” is not similar at all. Mr. Gupta’s own Exhibits A, H and I illustrate the point—even to a non-programmer, the Linux code is completely different from the Unix code—and detailed examination of the code shows no relationship.
22. Another two items identified by Mr. Gupta (the IPC header files and the ELF header file) are no more protectable. The IPC material is in the public domain, because, as I understand it, it was included in Unix System V 2.0 and Unix System V 3.2 without copyright notices and distributed before 1989. Moreover, to the extent it is even original, the IPC material constitutes ideas, concepts or principles, merger material and scènes à faire material. Furthermore, in places, Mr. Gupta’s conclusions of similarity depend on his selecting isolated lines of code from disparate places and putting them



---

together as if contiguous blocks of code were involved (which they are not) and important differences did not exist (which they do).

23. The ELF header file is classic scènes à faire material. The contents of the file are determined by software standards, target industry practice and demands and computer industry programming practices. Indeed, the substantive content in the ELF header file comes from a published and widely distributed standard — the “Tool Interface Standard (TIS) Executable and Linking Format (ELF) Specification”, version 1.2, attached as Exhibit XVI— material that explicitly grants permission for use in the interests of interoperability, by an industry consortium that included SCO’s alleged predecessor-in-interest, The Santa Cruz Operation, Inc.
24. Mr. Gupta’s claims for “certain UNIX System V headers and interfaces” (¶¶ 63-72) and “SYS V init code” (¶¶ 73-76) concern code that is not found in the Linux kernel.<sup>1</sup> That code is therefore outside what I understand to be the scope of the IBM Counterclaim at issue. Mr. Gupta’s claims as to this code nevertheless suffer from the defects described above: a failure to filter non-protectable material and an inappropriate standard of similarity.
25. Each of these six regions involves only elements that are unprotectable and/or dissimilar, as is demonstrated in detail in Exhibits II through VII.

---


<sup>1</sup> The “kernel” is the fundamental part of an operating system. It is the piece of software responsible for providing secure access to the machine’s hardware. Since there are many programs, and access to the hardware is limited, the kernel is also responsible for deciding when and how long a program should be able to make use of a piece of hardware. [from WordIQ Dictionary, [www.wordiq.com](http://www.wordiq.com)]

26. Putting aside the errors in Mr. Gupta's methodology, the code he identifies as infringing does not constitute a substantial part of the allegedly copyrighted works. SCO purports to hold copyrights in several versions of Unix System V source code, which over the years have ranged in size from several hundred thousand lines of code (in early versions) to many millions of lines of code in a current version. The alleged similarities identified by Mr. Gupta amount to a total of less than three hundred lines.
27. The allegedly infringed portions of the allegedly copyrighted works are not significant or important parts of the allegedly copyrighted works, considered as a whole — either quantitatively or qualitatively.

**V. Conclusions**

28. Mr. Gupta's declaration contains errors of methodology and fact. His conclusions of substantial similarity are flawed because he fails to exclude from comparison unprotectable elements of the allegedly copyrighted code and he uses an indefensible standard for what qualifies as "substantially similar" code.
29. If unprotectable elements are excluded from the comparison and an appropriate standard of similarity is applied there is no similarity between the six parts of Linux identified by Mr. Gupta and the allegedly copyrighted works.
30. Errors in methodology aside, the code Mr. Gupta alleges to be infringing does not constitute a substantial part of the allegedly copyrighted works, amounting to less than three hundred lines, which pales in comparison to the millions of lines in a current version of Unix.

31. I declare under penalty of perjury that the foregoing is true and correct.



Brian W. Kernighan

8/20/04

Date

Princeton, NJ

Place

---

**Materials Considered During Preparation of this Declaration**

*Computer Associates v. Altai*, 982 F.2d 693 (2d Cir. 1992)

*Gates Rubber, Inc., v. Bando American, Inc.*, 9 F.3d 823 (10th Cir. 1993)

*Mitel, Inc. v. Iqtel, Inc.* 124 F.3d 1366 (10th Cir. 1997)

*System Calls (Solaris 2.4 manual section 2)*. SUN Microsystems, Inc., 1994.

*The Single UNIX Specification, Version 2*. The Open Group, 1997.  
<http://www.opengroup.org/onlinepubs/007908799/toc.htm>.

*The Single UNIX Specification, Version 3*. The Open Group, 2004.  
<http://www.opengroup.org/onlinepubs/009695399/nframe.html>.

Maurice J. Bach, *The Design of the UNIX Operating System*. Prentice-Hall, 1986.

W. Richard Stevens, *Advanced Programming in the UNIX Environment*. Addison-Wesley, 1992.

Éric Lévénez, *Unix Timeline*. <http://www.levenez.com/unix>.

Code cited by Gupta for both Unix and Linux

All exhibits to Gupta declaration

Web site [lxr.linux.no](http://lxr.linux.no) (Linux Cross-reference) for Linux source code

Code produced by SCO for Unix source code