

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
28 July 2005 (28.07.2005)

PCT

(10) International Publication Number
WO 2005/067430 A2

(51) International Patent Classification: **Not classified**

(21) International Application Number:
PCT/US2004/024026

(22) International Filing Date: 26 July 2004 (26.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/749,959 31 December 2003 (31.12.2003) US

(71) Applicant (for all designated States except US): **MICROSOFT CORPORATION** [US/US]; One Microsoft Way, Redmond, WA 98052 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MOHAMED, Ahmed, H.** [US/US]; 29 210th Place SE, Sammamish, WA 98074 (US). **VOELLM, Anthony, F.** [US/US]; 9800 229th Lane NE, Redmond, WA 98053 (US).

(74) Agent: **CONKLIN, John, B.**; Leydig, Voit & Mayer, LTD., Two Prudential Plaza, Suite 4900, 180 N Stetson Av, Chicago, IL 60601-6780 (US).

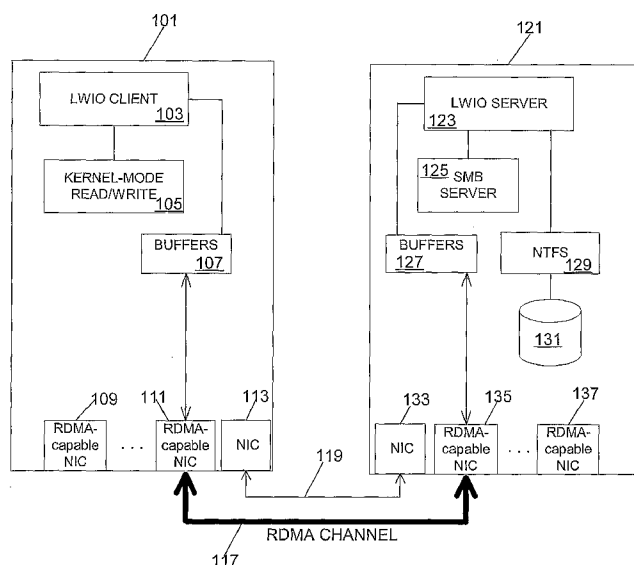
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: **LIGHTWEIGHT INPUT/OUTPUT PROTOCOL**



(57) Abstract: A method and system for offloading I/O processing from a first computer to a second computer, using RDMA-capable network interconnects, are disclosed. The method and system include a client on the first computer communicating over an RDMA connection to a server on the second computer by way of a lightweight input/output (LWIO) protocol. The protocol generally comprises a network discovery phase followed by an I/O processing phase. During the discovery phase, the client and server determine a minimal list of shared RDMA-capable providers. During the I/O processing phase, the client posts I/O requests for offloading to the second machine over a mutually-authenticated RDMA channel. The I/O model is asymmetric, with read operations being implemented using RDMA and write operations being implemented using normal sends. Read and write requests may be completed in polling mode and in interrupt mode. Buffers are managed by way of a credit mechanism.

WO 2005/067430 A2



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.