

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
18 December 2008 (18.12.2008)

PCT

(10) International Publication Number  
**WO 2008/154306 A1**

(51) International Patent Classification:

G06Q 40/00 (2006.01)

(21) International Application Number:

PCT/US2008/065955

(22) International Filing Date: 5 June 2008 (05.06.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

11/760,211 8 June 2007 (08.06.2007) US

(71) Applicant (for all designated States except US): **EXEGY INCORPORATED** [US/US]; 349 Marshall Avenue, Suite 100, St. Louis, MO 63119 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SINGLA, Naveen** [IN/US]; 530 Union Blvd., Apt. 605, St. Louis, MO 63108 (IN). **PARSONS, Scott** [US/US]; 10 Beaufort Court, St. Charles, MO 63301 (US). **FRANKLIN, Mark, A.** [US/US]; 7456 Stratford Avenue, St. Louis, MO 63130

(US). **TAYLOR, David, E.** [US/US]; 3448 Missouri Avenue, St. Louis, MO 63118 (US).

(74) Agent: **VOLK, Benjamin, L., Jr.**; Thompson Coburn LLP, One US Bank Plaza, St. Louis, MO 63101 (US).

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

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR HIGH SPEED OPTIONS PRICING

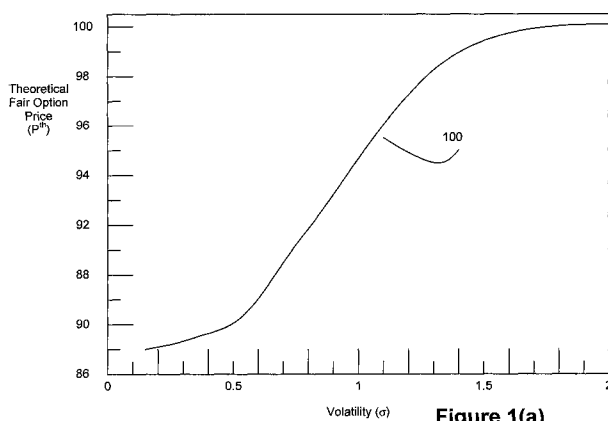


Figure 1(a)

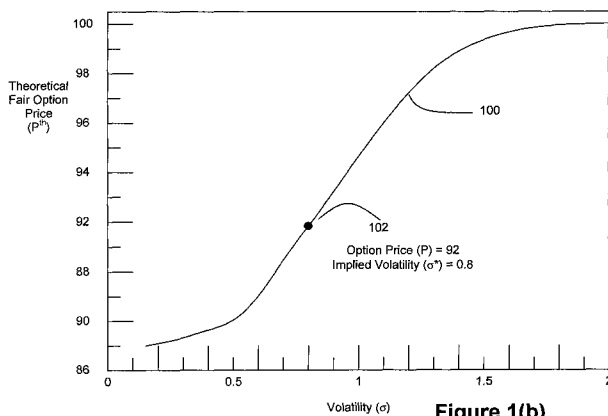


Figure 1(b)

(57) Abstract: A high speed technique for options pricing in the financial industry is disclosed that can provide both high throughput and low latency. A parallel/pipelined architecture is disclosed for computing an implied volatility in connection with an option. Parallel/pipelined architectures are also disclosed for computing an option's theoretical fair price. Preferably these parallel/pipelined architectures are deployed in hardware, and more preferably reconfigurable logic such as Field Programmable Gate Arrays (FPGAs) to accelerate the options pricing operations relative to conventional software-based options pricing operations.

WO 2008/154306 A1



FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL,  
NO, 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:**

— *with international search report*

Method and System for High Speed Options PricingCross-Reference and Priority Claim to Related Patent Applications:

5        This patent application claims priority to provisional U.S. patent application 60/814,796, filed June 19, 2006, and entitled "High Speed Processing of Financial Information Using FPGA Devices", the entire disclosure of which is incorporated herein by reference.

10        This patent application is related to the following patent applications: U.S. patent application 09/545,472 (filed April 7, 2000, and entitled "Associative Database Scanning and Information Retrieval", now U.S. patent 6,711,558), U.S. patent application 10/153,151 (filed May 21, 2002, and entitled "Associative Database Scanning and Information Retrieval using FPGA Devices", now U.S. patent 7,139,743), U.S. patent application 11/561,615 (filed November 20, 2006, entitled "Method and Apparatus for Processing Financial Information at Hardware Speeds Using FPGA Devices", and published as 2007/0078837), published PCT applications WO 05/048134 and WO 05/026925 (both filed May 21, 2004, and entitled "Intelligent Data Storage and Processing Using FPGA Devices"), U.S. provisional patent application 60/658,418 (filed March 3, 2005, and entitled "Biosequence Similarity Searching Using FPGA Devices"), U.S. provisional patent application 60/736,081 (filed November 11, 2005, and entitled "Method and Apparatus for Performing Biosequence Similarity Searching"), PCT patent application PCT/US2006/006105 (filed February 22, 2006, and entitled Method and Apparatus for Performing Biosequence Similarity Searching), U.S. patent application 11/293,619 (filed December 2, 2005, and entitled "Method and Device for High Performance Regular Expression Pattern Matching"), U.S. patent application 11/339,892 (filed January 26, 2006, and entitled "Firmware Socket Module for FPGA-Based Pipeline Processing"), and U.S. patent application 11/381,214 (filed May 2, 2006, and entitled "Method and Apparatus for Approximate Pattern Matching"), the entire disclosures of each of which are incorporated herein by reference.

Field of the Invention:

The present invention relates to the field of processing financial market data, particularly the field of performing options pricing operations on financial market data.

5

Background and Summary of the Invention:

Speed of information delivery is a valuable dimension to the financial instrument trading and brokerage industry. The ability of a trader to obtain analytical information (e.g., pricing, trade  
10 volume, trends, etc.) about financial instruments such as stocks, bonds and particularly options as quickly as possible cannot be overstated; reductions in information delivery delay on the order of fractions of a second can provide important value to traders. However, conventional techniques in the art which rely on software  
15 executed on a general purpose processor (GPP) to compute analytical information for financial instruments generally suffer latency issues due to the computation intensive nature of many financial instrument analytics models.

Options pricing is a function that particularly suffers from  
20 computational latency in a conventional GPP-based platform. An "option" is a derivative financial instrument that is related to an underlying financial instrument. An option is a contract that allows the option holder to trade in shares of the underlying financial instrument at a specific price at some time in the future. This  
25 future time is related to the option's lifetime (also referred to as the option's "time to maturity"), which is the amount of time after which the option to buy/sell cannot be exercised. The specific time in the future that the option-holder can exercise his/her choice depends on the type of the option; e.g., an American option can be  
30 exercised at any time during the option's lifetime while a European option can only be exercised at the end of the option's lifetime and yet other classes of options can be exercised only at certain pre-determined times during their lifetime.

In the parlance of the financial market, an option to buy  
35 shares is referred to as a "call option" (or "call" for short), and an option to sell shares is referred to as a "put option" (or "put"

for short). The inventors herein note that the teachings of the present invention are equally applicable to both call options and put options.

Within a financial market data stream, an offer to buy or sell an option is defined by the following characteristics: the identity of the financial instrument underlying the option (e.g., IBM stock), the number of shares of the underlying financial instrument that are covered by the option, the purchase price for the option ( $P$ ), the lifetime for the option ( $T$ ), the current price of the underlying financial instrument ( $S$ ), and the fixed price ( $K$ ) at which the option-holder has the choice to buy or sell the shares of the underlying financial instrument in the future (which is also known as the "strike price"). An option may additionally include a dividend yield ( $\delta$ ). Another parameter to be included in the trading of options, not directly related to the option but to the market, is the risk-free interest rate ( $r$ ) at which the potential option-holder can borrow or lend money. The modifier risk-free is used to signify that this is the rate of return one could expect on an investment that has zero risk involved with it. Further options characteristics that affect option pricing include whether the option is a call or put and whether the option is an American or European option, as explained above.

Based on knowledge of the values for these option characteristics and the risk-free interest rate, a trader may then assess whether the purchase price  $P$  for the option represents a good deal for the trader. While what constitutes a "good deal" will almost assuredly vary for different traders, it is expected that most traders will have an interest in buying or selling an option if the price  $P$  for the option is deemed favorable in view of the other option characteristics. Among the challenges in making this assessment as to an option's desirability is the complexity of factors that affect option price and option desirability.

One important consideration in making such an assessment of an option is an estimate of how much the price of the underlying financial instrument can fluctuate in the future. This fluctuation in the price of the underlying financial instrument is commonly

referred to as the "volatility" of the financial instrument. One measure of this volatility is to observe the historical trend of the price of the financial instrument and extrapolate the volatility from the historical trend. Another measure of the volatility can be obtained by approximating (within some degree of tolerance) the purchase price for the option (P) with a theoretical fair market option price ( $P^{\text{th}}$ ) that depends on the volatility. The volatility so obtained is termed the "implied volatility" since it is representative of the volatility of the underlying financial instrument implied by the option. Such a theoretical fair market option price ( $P^{\text{th}}$ ) can be calculated using an option pricing model. Various option pricing models are known in the art or used in a proprietary manner for computing theoretical option prices. Prominent among these models are the Black-Scholes option pricing model and the Cox, Ross, and Rubinstein (CRR) option pricing model, both of which are well-known in the art.

As indicated above, the downside of using some of these option pricing models to evaluate whether a buy/sell offer for an option represents a desirable transaction is that these option pricing models are very computation intensive. Because of their computation intensive nature, a delay is introduced between the time that the data regarding the option reaches a trader and the time when the trader has access to the pricing information computed from the option data (e.g., the option's implied volatility and/or the option's theoretical fair market price). This delay may be costly to the trader in that another party may have already bought the offered option while the trader was awaiting the results of the option pricing model. For example, consider the following exemplary scenario. Suppose there is an outstanding "bid" on an option for stock X that is a firm quote to sell an option to buy 100 shares of Stock X at an option price (P) of \$10.00 per share, and wherein the current market price (S) for Stock X is \$25 per share, and wherein the option further has values defined for K, T, and  $\delta$ , and wherein r is known. Also suppose there are two traders, A and B, each wanting to buy a call option on 100 shares of stock X, but before doing so would like to know whether the option price P of \$10.00 per

share represents a good deal given the other option characteristics ( $S$ ,  $K$ ,  $T$ , and  $\delta$ ) and the risk-free interest rate ( $r$ ). To aid this evaluation process, Trader A and Trader B both have their own implementations of an option pricing model; and they each would typically apply the values of these option characteristics to their respective implementations of an option pricing model to obtain information that is indicative of whether the call option offer represents a desirable transaction. For the purposes of this example, it will be presumed that this offer to sell the call option does in fact represent a good deal to both traders. The trader whose implementation of an option pricing model operates the fastest will have a decided market and investment advantage as he will be able to make an informed investment decision to purchase the call option before the other trader since the "winning" trader will be informed of the option's desirability while the "losing" trader is still awaiting the output from his/her respective implementation of the option pricing model. Accordingly, the "losing" trader will miss out on the desirable option because the "winning" trader will have taken that option offer off the market before the "losing" trader realized that he/she wanted to buy it. Thus, it can be seen that the speed of a trader's option pricing engine inevitably translates into a trading advantage, which even in a single large volume opportunity can amount to significant sums of money. Over time this market advantage can even lead to the success or failure of a trader to attract and keep customers and stay in business.

The ability of a computational engine that implements an option pricing model to quickly produce its output is even more significant when "black box" trading is taken into consideration. With such black box trading, the trader does not eyeball offers to buy/sell financial instruments as they tick across a trading screen to decide whether or not he/she will buy/sell a financial instrument. Instead, the trader defines the conditions under which he/she will buy/sell various financial instruments via a computer implemented algorithm. This algorithm then traverses the offers to buy/sell various financial instruments within a market data feed to identify which offers meet the specified conditions. Upon finding a

"hit" on an offer within the feed, the algorithm operates to automatically execute a specified trade on the offer (without further trader intervention). Thus, returning to the above example, such an algorithm may have a specified condition to the effect of

5 "buy a call option on X shares of ABC stock if the implied volatility for the call option is less than or equal to Z". Another exemplary algorithmic condition could be "buy a call option on Y shares of ABC stock if the computed theoretical fair market price for that option is greater than the actual price for the option by

10 at least Z cents". Thus, before the algorithm can make a decision as to whether a given call option offer will be purchased, the implied volatility and/or theoretical fair market price for the option offer will need to be computed via some form of an option pricing model. As explained above, with such black box trading, the

15 computation latency for computing the implied volatility and/or theoretical fair market price is highly important as delays on the order of fractions of a second will be critical in determining which trader is able to strike first to buy or sell options at a desirable price. Given that the inventors envision that black box trading

20 will continue to grow in prominence in future years (for example, it is estimated that currently greater than 50% of trades are performed automatically via computer-generated "black box" transactions), it is believed that high performance computation engines for option pricing will become ever more important to the financial instrument

25 trading industry.

Use of the CRR option pricing model for the analytics discussed above is especially computation intensive, as it is both an iterative model and a binomial model. Using the CRR option pricing model, an option's theoretical fair market price ( $P^{\text{th}}$ ) can be

30 computed as a function of the following inputs:  $P$ ,  $S$ ,  $K$ ,  $T$ ,  $\delta$ ,  $r$ , a volatility value  $\sigma$ , and  $n$ , wherein  $n$  represents a number of discrete time steps within the option's lifetime that the underlying financial instrument's price may fluctuate. The parameter  $n$  is specified by the user of the CRR option pricing algorithm. By

35 iteratively updating the volatility value  $\sigma$  until the theoretical fair market price  $P^{\text{th}}$  for that option approaches the option's actual

market price ( $P$ ), an option's "implied volatility" ( $\sigma^*$ ) can be computed. The "implied volatility", which represents the volatility of the underlying financial instrument at which the option's theoretical fair market price ( $P^{\text{th}}$ ) is within some specified tolerance  $\epsilon$  of the option's actual purchase price ( $P$ ), is an important characteristic of an option that is used by traders to decide whether a given option should be bought or not. However, because of the iterative nature of the implied volatility computation and the binomial nature of the theoretical fair market price computation, the calculation of implied volatility using the CRR option pricing model is highly computation intensive, as noted above. Conventional implementations of the CRR option pricing model to compute implied volatility in software on GPPs are believed by the inventors to be unsatisfactory because of the processing delays experienced while computing the implied volatility.

Based on the foregoing, the inventors herein believe that a technical problem exists in the art with respect to how one can accelerate the speed by which option pricing models can be used to evaluate option prices.

As further background, the inventors note that, in an attempt to promptly deliver financial information to interested parties such as traders, a variety of market data platforms have been developed for the purpose of ostensible "real time" delivery of streaming bid, offer, and trade information for financial instruments to traders.

Figure 12 illustrates an exemplary platform that is currently known in the art and used by traders to support their trading activities, including options trading. As shown in Figure 12, the market data platform 1200 comprises a plurality of functional units 1202 that are configured to carry out data processing operations such as the ones depicted in units 1202 (including options pricing), whereby traders at workstations 1204 have access to financial data of interest and whereby trade information can be sent to various exchanges or other outside systems via output path 1212. The purpose and details of the functions performed by functional units 1202 are well-known in the art. A stream 1206 of financial data arrives at the system 1200 from an external source such as the

exchanges themselves (e.g., NYSE, NASDAQ, etc.) over private data communication lines or from extranet providers such as Savvis or BT Radians. The financial data source stream 1206 comprises a series of messages that individually represent a new offer to buy or sell a financial instrument, an indication of a completed sale of a financial instrument, notifications of corrections to previously-reported sales of a financial instrument, administrative messages related to such transactions, and the like. As used herein, a "financial instrument" refers to a contract representing equity ownership, debt or credit, typically in relation to a corporate or governmental entity, wherein the contract is saleable. Examples of "financial instruments" include stocks, bonds, options, commodities, currency traded on currency markets, etc. but would not include cash or checks in the sense of how those items are used outside financial trading markets (i.e., the purchase of groceries at a grocery store using cash or check would not be covered by the term "financial instrument" as used herein; similarly, the withdrawal of \$100 in cash from an Automatic Teller Machine using a debit card would not be covered by the term "financial instrument" as used herein).

Functional units 1202 of the system then operate on stream 1206 or data derived therefrom to carry out a variety of financial processing tasks. As used herein, the term "financial market data" refers to the data contained in or derived from a series of messages that individually represent a new offer to buy or sell a financial instrument, an indication of a completed sale of a financial instrument, notifications of corrections to previously-reported sales of a financial instrument, administrative messages related to such transactions, and the like. The term "financial market source data" refers to a feed of financial market data received directly from a data source such as an exchange itself or a third party provider (e.g., a Savvis or BT Radians provider). The term "financial market secondary data" refers to financial market data that has been derived from financial market source data, such as data produced by a feed compression operation, a feed handling operation, an option pricing operation, etc.

Because of the massive computations required to support such a platform, current implementations known to the inventors herein typically deploy these functions across a number of individual computer systems that are networked together, to thereby achieve the appropriate processing scale for information delivery to traders with an acceptable degree of latency. This distribution process involves partitioning a given function into multiple logical units and implementing each logical unit in software on its own computer system/server. The particular partitioning scheme that is used is dependent on the particular function and the nature of the data with which that function works. The inventors believe that a number of different partitioning schemes for market data platforms have been developed over the years. For large market data platforms, the scale of deployment across multiple computer systems and servers can be physically massive, often filling entire rooms with computer systems and servers, thereby contributing to expensive and complex purchasing, maintenance, and service issues.

This partitioning approach is shown by Figure 12 wherein each functional unit 1202 can be thought of as its own computer system or server. Buses 1208 and 1210 can be used to network different functional units 1202 together. For many functions, redundancy and scale can be provided by parallel computer systems/servers such as those shown in connection with options pricing and others. To the inventors' knowledge, these functions are deployed in software that is executed by the conventional GPPs resident on the computer systems/servers 1202. The nature of GPPs and software systems in the current state of the art known to the inventors herein imposes constraints that limit the performance of these functions. Performance is typically measured as some number of units of computational work that can be performed per unit time on a system (commonly called "throughput"), and the time required to perform each individual unit of computational work from start to finish (commonly called "latency" or delay). Also, because of the many physical machines required by system 1200, communication latencies are introduced into the data processing operations because of the

processing overhead involved in transmitting messages to and from different machines.

Despite the improvements to the industry that these systems have provided, the inventors herein believe that significant further  
5 improvements can be made. In doing so, the inventors herein disclose that the underlying technology disclosed in the related and incorporated patents and patent applications identified above can be harnessed in a novel and non-obvious way to fundamentally change the system architecture in which market data platforms are deployed.

10 In above-referenced related U.S. patent 7,139,743, it was first disclosed that reconfigurable logic, such as Field Programmable Gate Arrays (FPGAs), can be deployed to process streaming financial information at hardware speeds. As examples, the '743 patent disclosed the use of FPGAs to perform data reduction  
15 operations on streaming financial information, with specific examples of such data reduction operations being a minimum price function, a maximum price function, and a latest price function.

Since that time, the inventors herein have greatly expanded the scope of functionality for processing streams of financial  
20 information with reconfigurable logic.

In accordance with one embodiment of the invention described herein, options pricing can be performed at hardware speeds via reconfigurable logic deployed in hardware appliances to greatly accelerate the speed by which option pricing operations can be  
25 performed, thereby providing important competitive advantages to traders. Thus, in accordance with this embodiment of the invention, it is disclosed that the options pricing functionality 1202 that is performed in software on conventional platforms can be replaced with reconfigurable logic that is configured as an options pricing  
30 engine. Such an options pricing engine can perform a number of computations related to options to aid in the evaluation of whether a given option represents a desirable transaction. For example, such an options pricing engine can be configured to compute an implied volatility for an option or a theoretical fair market price  
35 for an option. The inventors further disclose that in addition to options pricing, other functions of a conventional market data

platform can be deployed in reconfigurable logic, thereby greatly consolidating the distributed nature of the conventional market data platform into fewer and much smaller appliances while still providing acceleration with respect to latency and throughput.

5       As used herein, the term "general-purpose processor" (or GPP) refers to a hardware device that fetches instructions and executes those instructions (for example, an Intel Xeon processor or an AMD Opteron processor). The term "reconfigurable logic" refers to any logic technology whose form and function can be significantly  
10 altered (i.e., reconfigured) in the field post-manufacture. This is to be contrasted with a GPP, whose function can change post-manufacture, but whose form is fixed at manufacture. The term "software" will refer to data processing functionality that is deployed on a GPP. The term "firmware" will refer to data  
15 processing functionality that is deployed on reconfigurable logic.

According to another embodiment of the invention, the inventors herein have streamlined the manner in which an option's implied volatility and fair market price can be computed, thereby providing acceleration independently of whether such functionality  
20 is deployed in hardware or software. While it is preferred that the implied volatility and fair market price computations disclosed herein be performed via firmware pipelines deployed in reconfigurable logic, the inventors herein further note that the architectural and technical improvements with respect to how the  
25 implied volatility and/or fair market prices can be computed can also provide acceleration when performed in hardware on custom Application Specific Integrated Circuits (ASICs), in software on other platforms such as superscalar processors, multi-core processors, graphics processor units (GPUs), physical processor  
30 units (PPUs), chip multi-processors, and GPPs, or in a hybrid system involving exploitation of hardware, including reconfigurable hardware, and software techniques executing on a variety of platforms.

With respect to computing an option's implied volatility, the  
35 inventors disclose an exemplary technical solution wherein an iterative banded m-ary search within the option's volatility space

can be performed to identify the volatility value for which the option's theoretical fair market price, computed according to an option pricing model, approximates the option's actual purchase price to within a predetermined tolerance. This identified  
5 volatility value for which the option's computed theoretical fair market price approximates the option's actual purchase price to within a predetermined tolerance can then be used as the option's implied volatility.

With this iterative banded m-ary approach, a plurality m+1  
10 theoretical fair market prices are preferably computed in parallel for different volatility values within the volatility space for a given iteration, thereby providing acceleration with respect to the computation of the implied volatility. At least one, and preferably a plurality of conditions are tested to determine whether an  
15 additional iteration is needed to find the implied volatility. As explained hereinafter, a theoretical fair market price convergence property ( $\epsilon$ ) is preferably used as one of these conditions. Also as explained hereinafter, a volatility convergence property ( $\epsilon_\sigma$ ) is preferably used as another of these conditions.

20 Preferably, the option pricing model that is used to compute the option's theoretical fair market price is the CRR option pricing model. To accelerate the computation of the option's theoretical fair market price according to the CRR model, disclosed herein is an exemplary technical solution for parallelizing and pipelining the  
25 computation of the intermediate prices at different time steps within the CRR binomial tree, thereby providing acceleration with respect to the computation of the theoretical fair market price according to the CRR option pricing model.

In accordance with another embodiment of the invention,  
30 disclosed herein is another technical solution employing a lookup table to retrieve precomputed terms that are used in the computation of a European option's theoretical fair market price. Optionally, the theoretical fair market option price can then be used to drive a computation of the implied volatility as described above.

35 Preferably, this lookup table is indexed by the European option's volatility and time to maturity. Furthermore, in an embodiment

wherein the lookup table is used to compute a European option's theoretical fair market price but not its implied volatility, it is preferred that an additional lookup table of volatility values indexed by financial instruments be employed to identify a  
5 volatility value applicable to the underlying financial instrument of the subject European option. The lookup table terms retrieved from the table and indexed by the option's volatility and time to maturity can be fed to a combinatorial logic stage that is configured to accelerate the computation of the theoretical fair  
10 market price by parallelizing the computation of the constituent components of a summation formula for determining the option's theoretical fair market price.

In accordance with yet another embodiment of the invention, disclosed herein is another technical solution for directly  
15 computing the terms that are used in the computation of the option's theoretical fair market price. With this technique, a plurality of parallel computation modules are preferably employed to compute each term in parallel, thereby accelerating the overall computation of the option's theoretical fair market price.

20 Further still, to better map these computational modules onto available processing resources, a partitioning scheme can optionally be employed to distribute portions of the term computations across different processing resources.

As noted above, these parallel/pipelined architectures for  
25 computing an option's implied volatility and/or theoretical fair market price are preferably deployed in reconfigurable logic, thereby providing not only data processing at hardware speeds but also providing flexibility with respect to the parameters and models used in the computations. Preferably, a firmware pipeline is  
30 deployed on the reconfigurable logic to accelerate at least a portion of these parallel/pipelined architectures, as explained in greater detail hereinafter. However, as stated above, it should be noted that processing resources other than reconfigurable logic can be used to implement the streamlined options pricing architectures  
35 described herein.

These and other features and advantages of the present invention will be understood by those having ordinary skill in the art upon review of the description and figures hereinafter.

5 Brief Description of the Drawings:

Figures 1(a) and (b) depict an exemplary plot of a theoretical fair market price for an option calculated using the CRR option pricing model versus the option's volatility;

10 Figures 1(c)-(e) depict an iterative banded approach to the computation of an option's implied volatility for the plot of Figures 1(a) and (b);

Figure 2(a) depicts an exemplary flowchart for the iterative banded approach for computing an option's implied volatility;

15 Figure 2(b) depicts an exemplary embodiment of a computational pipeline for computing an option's implied volatility;

Figure 2(c) depicts an exemplary option message that can be processed by the present invention;

Figure 3 depicts an exemplary embodiment of a computational pipeline for an iterative stage of the pipeline of Figure 2(b);

20 Figure 4 depicts an exemplary embodiment of a binomial tree of depth n for computing a theoretical fair market option price based on the CRR option pricing model;

25 Figure 5(a) depicts an exemplary embodiment of a computational binomial tree for computing a theoretical fair market option price based on the CRR option pricing model;

Figures 5(b)-(d) depict an exemplary embodiment of combinatorial logic stage for computing a theoretical fair market option price based on the CRR option pricing model;

30 Figures 6(a) and (b) depict an exemplary technique for computing a European option's fair market price using lookup tables;

Figure 6(c) depicts an exemplary embodiment of the combinatorial logic stage of Figure 6(a);

Figure 7 depicts an exemplary embodiment for directly computing an option's fair market price;

35 Figures 8(a)-(c) depict exemplary embodiments of binomial trees for a CRR option pricing model applied to European options,

wherein the depth  $n$  of the tree is 4 and the number of operation needed to compute terminal option prices is minimized;

Figure 8(d) depicts an exemplary embodiment of a computational pipeline for computing an option's fair market price for the

5 binomial tree of Figure 8(b);

Figure 8(e) depicts an exemplary embodiment of a computational pipeline for computing an option's fair market price for the binomial tree of Figure 8(c);

10 Figure 9(a) depicts an exemplary embodiment of a binomial tree wherein partitioning is used;

Figures 9(b)-(d) depict exemplary embodiments of partitioned computational pipelines for the computation of an option's fair market price for the binomial tree of Figure 9(a);

15 Figure 10(a) depicts an exemplary embodiment of a binomial tree wherein double partitioning is used;

Figure 10(b) depicts an exemplary embodiment of a double partitioned computational pipeline for the computation of an option's fair market price for the binomial tree of Figure 10(a);

20 Figure 11 depicts an exemplary embodiment of a computational pipeline for implementing the computational module 702 of Figure 7;

Figure 12 depicts an exemplary system architecture for a conventional market data platform;

25 Figure 13 depicts an exemplary architecture for a market data platform wherein at least portions of the functional units are deployed in hardware;

Figure 14 is a block diagram view of an exemplary system architecture in accordance with an embodiment of the present invention on which an options pricing engine can be deployed;

30 Figure 15 illustrates an exemplary framework for the deployment of software and firmware for an embodiment of the present invention;

Figure 16(a) is a block diagram view of a preferred printed circuit board for installation into a market data platform to carry out data processing tasks in accordance with an embodiment of the  
35 present invention;

Figure 16(b) is a block diagram view of an alternate printed circuit board for installation into a market data platform to carry out data processing tasks in accordance with an embodiment of the present invention;

5        Figure 17 illustrates an example of how the firmware application modules of a pipeline can be deployed across multiple FPGAs;

10        Figure 18 illustrates an exemplary firmware application module pipeline for message processing wherein options pricing can be performed;

      Figure 19 illustrates another exemplary firmware application module pipeline for message processing wherein options pricing can be performed; and

15        Figures 20(a) and (b) depict an exemplary embodiment for implementing the OPM computational unit 300 of Figure 3 using a lookup table.

#### Detailed Description of the Preferred Embodiments:

20        In accordance with an embodiment of the present invention, it is desired to efficiently compute an option's implied volatility  $\sigma^*$ . Figure 1(a) is an exemplary plot 100 that illustrates how an option's theoretical fair market option price, computed using the CRR option pricing model, varies as a function of volatility. Theoretical fair market price refers to the price that should be  
25        assigned to the option to prevent riskless arbitrage opportunities, i.e., the option should be priced so that the potential option-holder cannot make a profit on it without taking any risk. Any of a number of option pricing models can be used to determine an option's theoretical fair market price, including but not limited to the CRR  
30        option pricing model, the Black-Scholes option pricing model, and other option pricing models, both those well-known to the public and those that are proprietary to traders and brokerage businesses. In this exemplary embodiment of the invention, the CRR option pricing model will be used; however it should be noted that a practitioner  
35        of the present invention can choose to deploy other option pricing models when computing an option's implied volatility.

Also, in this example, it is assumed that the stock price  $S$  for the stock underlying the option is 100, the strike price  $K$  for the option is 100, the time to maturity  $T$  for the option is 10 months, the interest rate  $r$  applicable to the option is 10%,  
5 compounded annually, the dividend yield  $\delta$  for the option is 0, and the number of time steps  $n$  for the CRR option pricing model is 100. It should be noted that the scale and values shown in this plot are exemplary only, as other values and scales may readily be used. For example, volatility is typically expressed in terms of standard  
10 deviation of the underlying financial instrument price around a mean value for the underlying financial instrument price. With such a scale, the volatility values would, by definition, always be positive. Given that the range of volatility values for most financial instruments will not exceed 2 standard deviations, for the  
15 purposes of explanation herein, it is helpful to illustrate a volatility range of 0 to 2. However, it should be noted that other ranges and scales could be used to express a financial instrument's volatility space.

Plot 100 shows that the option price is a monotonically  
20 increasing function of volatility. As used herein, monotonically increasing refers to a function wherein the first derivative thereof is always a non-negative number. With plot 100, as the volatility increases, the curve approaches the value  $S$  of the stock price (which has been assumed to be 100 in this example).

25 The problem of interest then is that given a purchase price for an option, one needs to calculate the volatility value for the option at that purchase price. It should be recalled that the implied volatility  $\sigma^*$  is the volatility value for which the option's theoretical fair market price approximates the option's actual  
30 purchase price  $P$  to within some tolerance. In this example, we will assume that the option's actual purchase price  $P$  is 92, which results in the option's implied volatility  $\sigma^*$  being 0.8, as shown by point 102 on plot 100 of Figure 1(b). Because the dependence of the theoretical fair market option price on volatility does not have a  
35 closed-form, an analytic expression cannot readily be used to directly compute the volatility as a function of the theoretical

fair market option price. One solution to this problem is an iterative method that iteratively computes the theoretical fair market option price from a given volatility value and iteratively adjusts the volatility value used in the computation until a  
5 theoretical fair market option price is computed that is within some tolerance of the actual option purchase price  $P$ .

To speed this process, parallelism is preferably applied, preferably in a banded iterative approach. With the banded iterative approach, multiple starting points for volatility values  
10 are chosen such that the volatility range is subdivided into a plurality  $m$  of bands, as shown in Figure 1(c). At each iteration, the theoretical fair market option price for each of a plurality of volatility values (preferably  $m+1$  volatility values ( $\sigma_1, \sigma_2, \dots, \sigma_{m+1}$ ) that define the boundaries of each volatility band) are computed.  
15 As noted, these computations are preferably performed in parallel. In the example of Figure 1(c), wherein  $m$  is 4, an initial volatility range of 0 to 2 has been subdivided into 4 bands, whose boundary volatility values are  $\sigma_1=0$ ,  $\sigma_2=0.5$ ,  $\sigma_3=1.0$ ,  $\sigma_4=1.5$  and  $\sigma_5=2.0$ . The theoretical fair market purchase prices ( $P^{th}(\sigma_1)$ ,  $P^{th}(\sigma_2)$ ,  $P^{th}(\sigma_3)$ ,  
20  $P^{th}(\sigma_4)$ , and  $P^{th}(\sigma_5)$ ) are computed for each of these volatility boundary values. After computation of these theoretical fair market option prices, the algorithm checks whether any of the computed theoretical fair market option prices are within some tolerance  $\varepsilon$  of the actual option purchase price  $P$ . If one of the theoretical  
25 purchase prices  $P^{th}(\sigma_k)$  is within  $\varepsilon$  of  $P$ , then the implied volatility  $\sigma^*$  is determined to be equal to  $\sigma_k$ . If none of the computed theoretical purchase prices are within  $\varepsilon$  of  $P$ , then another iteration of the algorithm is needed. Thus, the tolerance  $\varepsilon$  serves as a theoretical fair market price convergence property. As  
30 explained below, two other conditions - a iteration maximum and a volatility convergence property - can also optionally be tested to determine whether another iteration is needed.

To intelligently narrow the volatility band within which the algorithm searches for the implied volatility, the algorithm  
35 determines the volatility band within which the implied volatility

resides. Because the theoretical fair market option price is a monotonically increasing function of volatility, the volatility band within which the implied volatility resides can be quickly determined by identifying the volatility band for which the theoretical fair market option prices at its boundaries encompass the actual option purchase price  $P$ . In the example of Figure 1(c), the volatility band within which the implied volatility resides is labeled as 104. As can be seen, the volatility value at the lower boundary of band 104 is 0.5, for which the theoretical fair market option price is approximately 90, and the volatility value at the upper boundary of band 104 is 1.0, for which the theoretical fair market option price is approximately 95. Because the actual option purchase price of 92 falls between these two computed theoretical purchase price values, band 104 is identified as the volatility band within which the implied volatility resides.

Thus, during the next iteration, band 104 from the first iteration is subdivided into another plurality  $m$  of bands, as shown in Figure 1(d). Thus, in the example of Figure 1(d), the volatility band boundary values for the second iteration are  $\sigma_1=0.5$ ,  $\sigma_2=0.625$ ,  $\sigma_3=0.75$ ,  $\sigma_4=0.875$  and  $\sigma_5=1.0$ . The theoretical fair market purchase prices are then computed for each of these volatility boundary values, and a check is once again performed to assess whether the implied volatility has been found by determining whether any of the computed theoretical fair market option prices are within  $\epsilon$  of  $P$ .

If the implied volatility has not been found, then the band 106 within which the implied volatility resides is once again identified. The identified band 106, is then further subdivided into a plurality  $m$  of bands during a next iteration, as shown in Figure 1(e). During this next iteration, the parallel computation of theoretical fair market option prices is once again performed with a check to determine whether the implied volatility has been found. Presuming that it has not, the algorithm operates to identify band 108 for subdivision during the next iteration. Such iterative narrowing of the volatility band within which the implied volatility resides then continues until the algorithm determines that it has found the implied volatility within a sufficient degree

of precision or any one of the other conditions for halting the process are satisfied.

While the example of Figure 1(a)-(e) has used a value of  $m=4$  as an example, it should be noted that other values of  $m$  could readily be used in this iterative banded search for the implied volatility. Also, while the same value of  $m$  has been used for each iteration in the examples of Figures 1(a)-(e), it should be noted that the value of  $m$  could be different for each iteration if so desired by a practitioner of this embodiment of the present invention. Further still, in the examples of Figures 1(a)-(e), each volatility band has been equally sized by essentially dividing the volatility range of interest by  $m$ . It should be noted that the volatility bands used in each iteration need not be equally sized. Moreover, in this example, the initial volatility range was set equal to a range of 0-2 (or a full range of volatility). It should be noted that other initial volatility ranges can be used. For example, a practitioner of this embodiment of the invention may find it more efficient to intelligently select the initial volatility range. With such intelligent selection, the initial volatility range can be selected based on historical volatility values for the option's underlying financial instrument. For example, the lowest recorded volatility value for the financial instrument in the past 12 months (optionally minus some tolerance) can be used as the lower bound of the initial volatility space, and the highest recorded volatility value for the financial instrument in the past 12 months (optionally plus some tolerance) can be used as the upper bound of the initial volatility space.

Figure 2(a) illustrates an exemplary flowchart for an iterative banded search for implied volatility. At step 200, the algorithm receives a message regarding an option that is available on the market as well as other option pricing parameters. As shown in Figure 2(c), this option message 260 preferably identifies the financial instrument underlying the option (e.g., IBM stock) (field 262), the number of shares covered by the option, and the  $P$ ,  $S$ ,  $K$ ,  $T$ , and  $\delta$  values for the option (fields 264, 266, 272, 268, and 274 respectively). As shown in Figure 2(c), the option message 260 also

preferably includes a field 270 that identifies whether the option is a call or put (the C/P flag). It should be noted that the name of the option itself in field 262 encodes whether the option is a European option or an American option. As such, the question of whether a given option is European and American can be resolved by analysis of the name field 262. Examples of option pricing parameters that are also preferably received at step 200 include the risk-free interest rate  $r$  and the number of time steps  $n$  to be used in the option pricing model. At step 202, the algorithm initializes its iteration index  $q$  to  $q=1$ . Then, at step 204, the algorithm defines the volatility band boundary values  $\sigma^q_1, \sigma^q_2, \dots, \sigma^q_{m+1}$ , for the first iteration.

Next at step 206, the algorithm computes the theoretical fair market option price  $P^{th}$  for each of the volatility band boundary values defined at step 204. Preferably, these computations are performed in parallel as described hereinafter. At step 208, the algorithm computes the differences  $D^q_1, D^q_2, \dots, D^q_{m+1}$  between each computed theoretical fair market option price and the actual option purchase price  $P$ , wherein  $D^q_k = P^{th,q}(\sigma_k) - P$ . As with step 206, the computations of step 208 are preferably performed in parallel. Next, at step 210, the algorithm checks whether the absolute value of any of the computed differences  $D^q_1, D^q_2, \dots, D^q_{m+1}$  is less than  $\epsilon$ . If the absolute values of one of the computed differences  $D^q_k$  is less than  $\epsilon$ , then at step 212, the algorithm sets the implied volatility  $\sigma^*$  equal to the volatility value  $\sigma^q_k$  for which the computed  $D^q_k$  is less than  $\epsilon$ . The algorithm can then stop and output the determined implied volatility.

However, if none of the differences  $D^q_k$  are less than  $\epsilon$ , then the algorithm proceeds to steps 214 and 216. Steps 214 and 216 operate to identify the volatility band within which the implied volatility resides. At step 214, the algorithm computes the products  $D^q_k * D^q_{k+1}$  for  $k=1, 2, \dots, m+1$ . Because of the monotonically increasing characteristic of the theoretical fair market option price with respect to volatility, there will only be one product that is a negative number because there will only be one band where the actual purchase price falls between the band's theoretical fair

market option price boundary values. Thus, at step 216, the algorithm will identify the band within which the implied volatility resides by determining which product is a negative number. The volatility boundary values  $\sigma_k^q$  and  $\sigma_{k+1}^q$  for which the product of  $D_k^q * D_{k+1}^q$  is negative will then be used to define the volatility range to be subdivided during the next iteration (step 218).

Preferably, the algorithm also uses two other parameters to control the number of iterations that the algorithm will perform -  $Q_{\max}$  and  $\epsilon_\sigma$ .  $Q_{\max}$  defines the maximum number of iterations that the algorithm will perform, and  $\epsilon_\sigma$  describes a volatility convergence property that, when met, will cause the algorithm to stop and output an implied volatility value. The  $Q_{\max}$  condition operates to decrease the latency of computation by preventing the occurrence of an excessive number of iterations. The  $\epsilon_\sigma$  condition also operates to decrease computational latency in cases where the volatility band search space around the implied volatility is very small, but because of a steep slope of the theoretical fair market price curve within that volatility space, the theoretical fair market price convergence condition  $\epsilon$  has not been satisfied. Accordingly, the theoretical fair market prices for the different volatility values within the volatility band may possess differences in value that exceed  $\epsilon$  while the differences between the different volatility values may be virtually negligible. In such instances, the average volatility value of the narrow volatility band in question is preferably output as the implied volatility. To control the algorithm, using these two parameters, steps 220, 222 and 224 operate as follows. At step 220, the algorithm computes the difference  $|\sigma_k^q - \sigma_{k+1}^q|$  for the volatility boundary values defined at step 218, and then compares this difference with  $\epsilon_\sigma$ . If the difference  $|\sigma_k^q - \sigma_{k+1}^q|$  is less than  $\epsilon_\sigma$ , then the algorithm determines that the defined volatility boundaries and the implied volatility have sufficiently converged, and at step 222, the algorithm outputs the implied volatility as the average of  $\sigma_k^q$  and  $\sigma_{k+1}^q$  ( $\sigma^* = (\sigma_k^q + \sigma_{k+1}^q) / 2$ ). If volatility convergence has not been achieved, then the algorithm proceeds to step 224, wherein it checks whether the

current iteration  $q$  is equal to  $Q_{\max}$ . If  $q$  does equal  $Q_{\max}$ , then the algorithm returns to step 222 to output the implied volatility as the average of  $\sigma_k^q$  and  $\sigma_{k+1}^q$ . If  $q$  is less than  $Q_{\max}$ , then the algorithm increments the iteration index  $q$  at step 226. Thereafter,  
 5 the algorithm returns to step 204 for the next iteration, wherein the  $m+1$  volatility band boundary values are defined starting from the volatility range defined at step 218.

The algorithm of Figure 2(a) can be deployed in either hardware or software (or a combination thereof). If deployed in  
 10 software, preferably a plurality of parallel co-processors or a multi-core processor is used to parallelize and accelerate at least step 206. However, for greater acceleration, preferably at least a portion of the algorithm of Figure 2(a) (and more preferably the full algorithm of Figure 2(a)) is deployed in hardware. Preferably,  
 15 this hardware is reconfigurable logic, as explained hereinafter.

Figure 2(b) depicts a computational pipeline 250 for realizing the algorithm of Figure 2(a). The pipeline 250 preferably comprises a plurality of computational modules 252 that are sequenced together to form the pipeline 250. Each computational module 252 is  
 20 preferably configured to perform the computations for one iteration of the algorithm of Figure 2(a). As such, it is preferred that the number of computational modules 252 in the pipeline 250 be  $Q_{\max}$ . The first computational module 252 in the pipeline preferably receives the option message 260 as an input (e.g., an option on IBM stock  
 25 with various defined values for  $P$ ,  $S$ ,  $T$ ,  $K$  and  $\delta$ ). The computational module 252 for the first iteration is preferably also initialized with  $m+1$  volatility band boundary values ( $\sigma_{1,1}^1, \sigma_{2,1}^1, \dots, \sigma_{m+1,1}^1$ ). The computational module 252 for the first iteration is also preferably initialized with values for  $n$ ,  $r$ ,  $\varepsilon$  and  $\varepsilon_0$ . The  
 30 computational modules 252 are then configured to execute steps 204-222 of the algorithm. The output of the first computational module 252 for receipt by the computational module 252 of the second iteration would then comprise the same option message parameters, the volatility band boundary values for the second iteration as  
 35 defined by step 204, and the  $n$ ,  $r$ ,  $\varepsilon$  and  $\varepsilon_0$  parameters used during the first iteration. The computational modules 252 will also

preferably output an implied volatility value  $\sigma^*$  if applicable. Due to the pipelined nature of the computational modules 252, the computational module 252 for iteration 1 can be working on a given option message at the same time that computational module 252 for iteration 2 is working on another different option message, and so on. As such, a high throughput of option messages can be streamed through the pipeline, all the while decreasing the latency with which the implied volatility for each option message is determined.

Figure 3 provides an exploded view of a computational module 252. Preferably, each computational module comprises a plurality  $m+1$  of parallel option pricing model (OPM) computational units 300. In the example of Figure 3, it can be seen that  $m$  is 4; however it should be understood that other values of  $m$  could readily be used. Each OPM computational unit 300 is configured to compute a theoretical fair market option price for an input volatility value. The output of each OPM computational unit 300 is fed to a computational unit 302 that computes the difference value  $D_k^q$ . Thus, the plurality of parallel OPM computational units 300 and difference computational units 302 perform steps 206 and 208 for the algorithm of Figure 2(a). A combinatorial logic stage 304 of the computational module 252 then receives the outputs from all of the difference computational units 302 and the various parameters  $P$ ,  $S$ ,  $K$ ,  $T$ ,  $\delta$ ,  $n$ ,  $r$ ,  $\varepsilon$ ,  $\varepsilon_\sigma$ , and  $\sigma_1^q, \sigma_2^q, \dots, \sigma_{m+1}^q$  to perform steps 210-222 of the algorithm.

Each computational module 252 also preferably receives as an input a  $\sigma^*$ \_enable signal that indicates whether the implied volatility  $\sigma^*$  for that option message has been found as well as a  $\sigma^*$  signal that identifies the value of the determined implied volatility (although the computational module 252 for the first iteration need not receive these signals as inputs). Combinatorial logic stage 304 preferably sets the  $\sigma^*$ \_enable signal high if it has found the implied volatility for the option message and then outputs the implied volatility  $\sigma^*$ . Preferably, a control unit such as the firmware socket module 1420 described hereinafter with respect to a reconfigurable logic device embodiment of this aspect of the invention operates to read the  $\sigma^*$ \_enable and  $\sigma^*$  outputs from each

computational module 252 to halt the pipelined processing for any message whose implied volatility has already been found. With such control, when an option's implied volatility has been found at iteration 3, the pipeline 250 need not wait until the final  $Q_{\max}$  iteration of the pipeline before outputting the implied volatility for the option, thereby eliminating unnecessary latency. The combinatorial logic stage 304 also preferably outputs the  $P$ ,  $S$ ,  $K$ ,  $T$ ,  $\delta$ ,  $n$ ,  $r$ ,  $\varepsilon$ ,  $\varepsilon_{\sigma}$ , and  $\sigma^{q+1}_1, \sigma^{q+1}_2, \dots, \sigma^{q+1}_{m+1}$  parameters for use by the next downstream computational module 252 in the pipeline 250.

As noted above, any of a variety of OPMs can be used by the OPM computational units 300 to compute the option's theoretical fair market price. However, it is preferred that an OPM based on the CRR OPM be used. Cox, Ross and Rubinstein showed that in order to calculate the theoretical fair market price of an option, one needs to know only the strike price  $K$ , the price  $S$  of the financial instrument underlying the option, the range of movement of  $S$  (i.e., the volatility), the time to maturity  $T$  for the option, the interest rate  $r$ , and the dividend yield  $\delta$  for the option. The CRR binomial model divides the lifetime  $T$  of the option into  $n$  discrete steps, and assumes that at each step, the stock price  $S$  can either go up by a fixed multiplicative factor  $u$  or go down by a fixed multiplicative factor  $d$ . The factors  $u$  and  $d$  depend on the volatility  $\sigma$  of the stock, the time to maturity of the options, and the number of discrete steps  $n$ . The upward and downward movement factors can be calculated as follows:

$$u = e^{\sigma\sqrt{T/n}}$$

(1)

$$d = \frac{1}{u}$$

(2)

These particular forms for  $u$  and  $d$  are chosen to ensure that, as the number of time steps  $n$  approaches infinity, the movement of the stock price according to the binomial model approximates the distribution of stock prices that was proposed by Black and Scholes. It should be noted that there are other formulations for  $u$  and  $d$  for which the movement of the stock price according to the binomial

model approximates the distribution of stock prices that was proposed by Black and Scholes. Also, the binomial model preferably assumes that the interest rate  $r$  and the dividend yield  $\delta$  remain constant over the lifetime of the option.

5 Under the assumption that the stock price changes multiplicatively, a binomial tree can be constructed that, at each time step, represents the possible values of the stock price. Figure 4 illustrates such a binomial tree 400. Note that at time step  $i$ , the stock price can take one of  $i+1$  values  $u^j d^{i-j} S$ , for  $j =$   
 10  $0, 1, \dots, i$ . Thus, starting from the root node 402 of  $S$ , the nodes 404 at time step 1 of the tree 400 exhibit values of  $uS$  and  $dS$ . At time step 2 of tree 400, the leaves 406 exhibit values of  $u^2 S$ ,  $udS$  and  $d^2 S$ . At time step 3, the leaves 408 exhibit values of  $u^3 S$ ,  $u^2 dS$ ,  $ud^2 S$  and  $d^3 S$ , and so on until time step  $n$ , wherein the leaves 410  
 15 exhibit values of  $u^n S$ ,  $u^{n-1} dS$ ,  $u^{n-2} d^2 S \dots u^2 d^{n-2} S$ ,  $ud^{n-1} S$  and  $d^n S$ .

The theoretical fair market price of the option can then be obtained by working backward through tree 400 from the leaves 410 at time step  $n$  through the root 402. Let  $C(n, i, j)$  denote the price of the option at the  $j^{\text{th}}$  node at time step  $n-i$  for a binomial tree of  
 20 depth  $n$ . Then,  $C(n, i, j)$  can be calculated as:

$$C(n, i, j) = \max \left[ R(pC(n, i-1, j+1) + (1-p)C(n, i-1, j)), Su^j d^{i-j} - K \right] \quad (3)$$

wherein  $R$  represents the normalized interest rate, wherein

$$R = e^{-rT/n} \quad (4)$$

25 and wherein:

$$p = \frac{(e^{(r-\delta)T/n} - d)}{(u - d)} \quad (5)$$

The computation of  $C(n, i, j)$  is performed for  $j = 0, 1, \dots, n-i$ , and  
 30 then the computation moves on to time step  $n-i-1$ . The computation is initialized at time step  $n$  by noting that, at the expiration of the option, the payoff at node  $j$  at time step  $n$  is  $\max(0, u^j d^{n-j} S - K)$ . This then should be the price of the option at that node so that riskless profitable arbitrage is not possible. Thus, at time step  
 35  $n$ , the computations are initialized as follows:

$$C(n,0,j)=\max(0,u^j d^{n-j} S-K), \text{ for } j=0, 1, \dots, n$$

(6)

It should be noted that formulas other than formula (3) for  $C(n,i,j)$  can be used in the practice of this embodiment of the present invention, as would be understood by those having ordinary skill in the art upon review of the teachings herein.

It should also be noted that formula 3 is directed toward the theoretical fair market price of call options. For put options, the theoretical fair market price can be computed as

$$C(n,i,j)=\max\left[R(pC(n,i-1,j+1)+(1-p)C(n,i-1,j)), K-Su^j d^{|i-j|}\right]$$

(3a)

which initializes at time step  $n$  as follows:

$$C(n,0,j)=\max(0,K-u^j d^{n-j} S), \text{ for } j=0, 1, \dots, n$$

(6a)

Further still, it should be noted that formulas 3 and 3a can be used to compute the theoretical fair market prices for both American and European options.

Preferably the computations for  $C(n,i,j)$  are implemented in a parallel pipeline architecture, as shown in Figure 5(a). Figure 5(a) depicts an exploded view of the OPM computational unit 300 of Figure 3. OPM computational unit 300 can be realized via a plurality of pipeline stages 504, wherein each pipeline stage 504 comprises at least one computational node 502. Each pipeline stage is configured to perform the  $C(n,i,j)$  computations for a given time step. The computational nodes 502 in the first pipeline stage 504 are preferably initialized with the  $C(n,0,j)$  values as determined by formula (6) for call options and by formula (6a) for put options. Each computational node 502 also preferably receives as inputs the parameters  $\sigma$ ,  $S$ ,  $K$ ,  $T$ ,  $r$ ,  $n$ ,  $\delta$  as well as the value of the C/P flag that identifies whether the pertinent option is a call or a put. Each computational node 502 is configured to perform the computation defined by formula (3) for  $C(n,i,j)$  if the option is a call option and the computation defined by formula (3a) for  $C(n,i,j)$  if the option is a put option, as defined by the option's C/P flag. In a reconfigurable logic implementation of computational node 252, each

node 252 can include parallel computational paths - one path for formula (3) and the other path for formula (3a), wherein control logic can be deployed to route pertinent input values to the appropriate computational path based on the option's call/put status. As such, each computed value for  $C(n,i,j)$  is fed to two downstream computational nodes 502 in the next pipeline stage 504 for use in the computation defined by formulas (3) and (3a).

The parallel pipelined nature of the computations performed by the OPM computational unit 300 of Figure 5(a) significantly accelerates the option pricing computations relative to conventional designs. Assume that it takes  $f_1$  clock cycles to perform the computations at each leaf node 410 of the binomial tree 400 and  $f_i$  clock cycles to perform the computations at each internal node, including the root. For a tree 400 of depth  $n$ , there are  $n+1$  leaf nodes 410 and  $n(n+1)/2$  internal nodes (including the root). The total number of clock cycles to compute the option price on a sequential machine would be  $(n+1)f_1 + 0.5n(n+1)f_i$ . On the other hand, using a parallel architecture such as that shown in Figure 5(a), all of the computations at each time step can be performed concurrently; thus reducing the number of clock cycles to  $(n+1)f_1 + nf_i$ . The complexity for the parallel pipelined architecture then becomes linear in the depth of the tree, as opposed to being quadratic in the depth of the tree as it is for a sequential machine. This aspect of the Figure 5(a) embodiment provides practitioners of the present invention with improved flexibility when designing and using an option pricing engine. A practitioner can use larger values of  $n$  relative to what he/she could use in a sequential machine for more accurate option price modeling without sacrificing latency, or a practitioner can price an option at much higher speeds, i.e., achieve a higher throughput, by using the same or similar values for  $n$  that he/she would have used for a sequential machine.

Figures 5(b)-(d) depict an exemplary embodiment for the combinatorial logic stage 304 of Figure 3. Figure 5(b) depicts a portion of stage 304 that is configured to perform steps 210 and 212 from Figure 2(a). As shown in Figure 5(b),  $m+1$  parallel absolute

value units 510 operate to find the absolute values of each input value for  $D_i^q$ . The computed absolute values  $|D_i^q|$  are then fed to the  $m+1$  parallel comparators 512, whereat each value of  $|D_i^q|$  is compared with the theoretical fair market price convergence property

5  $\epsilon$ . It should be noted that the comparators 512 are preferably be configured to output a value of 1 if  $|D_i^q|$  satisfies the fair market price convergence property and output a value of 0 otherwise. If one of the differences  $|D_i^q|$  satisfies the  $\epsilon$  condition, then the corresponding multiplier 514 will multiply a value of 1 times  $\sigma_i^q$  to  
10 produce the implied volatility value  $\sigma^*$ , which in turn is summed by adder 516 with the other 0 values produced by the other multipliers. The OR gate 518 will operate to set the  $\sigma^*$ \_enable signal high by detecting when one of the multipliers 514 produces an output of 1. Multipliers 514, adder 516, and OR gate 518 thus operate to output  
15 appropriate values of  $\sigma^*$  and  $\sigma^*$ \_enable.

Figure 5(c) depicts a portion of stage 304 that is configured to perform steps 214, 216, 218, 220, 222 and 204 from Figure 2(a). A plurality  $m$  of parallel multipliers 520 operate to multiply each successive difference  $D_i^q$  and  $D_{i+1}^q$  with each other. Only one of  
20 these multiplication products will be negative because there will only be one volatility band where the option's actual purchase price is greater than the theoretical fair market price at one volatility band boundary and less than the theoretical fair market price at the other volatility boundary for that band. Comparators 522 then  
25 operate to identify which of the multipliers 520 produces a negative product. Preferably, the output  $S_i$  of each comparator 522 is configured to go high only if the input product  $D_i^q * D_{i+1}^q$  is negative. A logic stage 524 then operates to compute the boundary values  $\sigma_k^q$  and  $\sigma_{k+1}^q$  of the volatility band within which the implied volatility  
30 resides. Adder 526, absolute value operator 528 and comparator 530 operate to test the identified volatility band defined by  $\sigma_k^q$  and  $\sigma_{k+1}^q$  against the volatility convergence property  $\epsilon_\sigma$ . If the volatility convergence property is satisfied, then comparator 530 asserts  $\sigma^*$ \_enable. Adder 532 and multiplier 534 operate to average  
35 the identified volatility band boundaries together; this average

serves as the implied volatility if the volatility convergence property is satisfied. A logic stage 536 operates to compute the volatility boundary values  $\sigma^{q+1}_1, \sigma^{q+1}_2, \dots, \sigma^{q+1}_m$  for the next iteration. Depending on whether the volatility convergence property is

5 satisfied, multiplexer 542 passes either the implied volatility or next iteration volatility values as output 542. It should be noted that in the final iteration module 252 in pipeline 250, the combinatorial logic stage 304 need not include logic 536 or multiplexer 540 since there will not be a need to compute a next

10 iteration's volatility values. Similarly, the combinatorial logic stage 304 of the final iteration module 252 in pipeline need not test the identified volatility band boundaries against the volatility convergence property.

It should also be noted that the combinatorial logic stage 304

15 can include control logic to reconcile situations where both the theoretical fair market price convergence property defined by  $\epsilon$  and the volatility convergence property defined by  $\epsilon_\sigma$  are met. In such instances, appropriate control logic can be included to pass the desired implied volatility as an output. If it is preferred that

20 the implied volatility as determined by  $\epsilon$  be used, then the control logic can be configured to pass the  $\sigma^*$  value identified at step 212 as the output. If it is preferred that the implied volatility as determined by  $\epsilon_\sigma$  be used, then the control logic can be configured to pass the  $\sigma^*$  value identified at step 222 as the output.

25 Alternatively, the control logic can be configured to pass the average of the two  $\sigma^*$  values identified at steps 212 and 222 as the output. Further still, the control logic can be configured to pass one of the  $\sigma^*$  values as an output based on which tolerance ( $\epsilon$  or  $\epsilon_\sigma$ ) satisfies one or more conditions.

30 Figure 5(d) depicts the logic stages 524 and 536 in greater detail. Within logic stage 524, a plurality  $2m$  of multipliers 550 operate to multiply each  $S_i$  value by  $\sigma^q_i$  and  $\sigma^q_{i+1}$ . The output of adders 552 will be the identified volatility band boundary values  $\sigma^q_k$  and  $\sigma^q_{k+1}$ . Within logic stage 536, an inverter 554 and adder 556

35 operate to generate the sum  $\sigma^q_k - \sigma^q_{k+1}$ , which represents the width of

the identified volatility band. Multipliers 560 operate to divide this bandwidth into m subbands of equal width, and adders 562 compute the resultant intermediate volatility values for the identified band to be used during the next iteration. It should be noted that logic stage 536 need not include a separate inverter 554 and adder 556; if desired, stage 536 can tap into the output of adder 526 shown in Figure 5(c).

Thus, by determining an option's implied volatility via the parallelized pipelined architecture shown in Figures 2(b), 3 and 5(a)-(d), significant enhancements to latency and throughput can be provided, thereby providing practitioners of the present invention with an important competitive advantage in the financial marketplace.

It should be noted that the iterative banded m-ary search disclosed herein in connection with Figures 1(a)-5 has been described in the context of an option's theoretical fair market price that is a monotonically increasing function of the option's volatility. The inventors further note that the iterative banded m-ary search for a particular parameter of interest related to a financial instrument can be performed in connection with other plots where a first parameter related to a financial instrument is a monotonically increasing (or monotonically decreasing) function of a second parameter related to the financial instrument. Further still, the inventors note that the first financial parameter can also be a monotonically strictly increasing or a monotonically strictly decreasing function of the second financial parameter. As used herein, monotonically decreasing refers to a function wherein the first derivative thereof is always a non-positive number, monotonically strictly increasing refers to a function wherein the first derivative thereof is always a number greater than zero, and monotonically strictly decreasing refers to a function wherein the first derivative thereof is always a number less than zero.

In accordance with another embodiment of the present invention, it is desired to efficiently compute an option's theoretical fair market price. As previously noted, the quick computation of fair market prices for options is a high priority

problem for traders, particularly for traders who perform trades using black box trading algorithms. If a trader or a trader's black box algorithm can detect a difference in the offered price P of an option and that option's theoretical fair market price (according to some option pricing model), then a trading opportunity may exist for that trader.

Figure 6(a) illustrates an exemplary architecture for low latency computation of a theoretical fair market price 624 from an option message input 600. In the example of Figure 6(a), it will be assumed that the option message pertains to a European call option, and it will further be assumed that the option pricing model employed by the architecture of Figure 6(a) to compute a theoretical fair market price for the European call option according to the CRR model is as follows:

$$P^{th} = \sum_{j=0}^n p^j q^{n-j} b_j \max[Su^j d^{n-j} - K, 0]$$

(7)

However, it should be noted that option pricing models other than formula (7) could be employed in the practice of this embodiment of the present invention.

With formula (7), the binomial coefficient,  $b_j$ , denotes the number of occurrences of j up states in n time steps, wherein:

$$b_j = \binom{n}{j} = \frac{n!}{j!(n-j)!}$$

(8)

The variables u and d are the incremental up and down factors for the financial instrument price S for each time step as described by formulas (1) and (2) above. The variables p and q are interval prices for up states and down states, respectively, wherein:

$$p = \frac{R - d}{R(u - d)}$$

(9)

$$q = \frac{1}{R} - p$$

(10)

The variable R is the normalized interest rate, wherein:

$$R = e^{\frac{r \times T}{n}}$$

(11)

If one examines each term in the summation of the binomial pricing model of formula (7), it is observed that the up/down factors (u,d) and the up/down state prices (p,q) are solely dependent on the number of time steps n, the risk-free interest rate r, and the volatility  $\sigma$ . Next, for purposes of this example, one makes the following assumptions: (1) an assumption (following from the Black-Scholes model) that the risk-free interest rate is constant over the life of the option, which helps simplify the options pricing algorithm, (2) an assumption that the number of time steps n in the model is the same for all financial instruments, which ensures that the computational complexity (and thus the latency) is constant for pricing different options, and (3) an assumption that the options share the same maturity dates. As such, one will assume that 3 month call options will mature on the same day, regardless of the underlying financial instrument. While not all options that share the same duration will share the same maturity date, one can further assume that the number of future maturity dates will be reasonably small (on the order of 10). This assumption is a reasonable one for European options since they can only be exercised at maturity; thus options maturing in a particular month will have very similar time to maturities, provided they all had a similar start date.

With the architecture 600 shown in Figures 6(a) and (b), the values of  $p^j q^{j-n} b_j$  and  $u^j d^{n-j}$  are precomputed for all values of j and stored in a lookup table 610 that is indexed by volatility  $\sigma$  and by time to maturity T. Thus,  $A(\sigma, T)$  will represent the array of n+1 values of  $p^j q^{j-n} b_j$  for  $0 \leq j \leq n$  and for T, and  $B(\sigma, T)$  will represent the array of n+1 values of  $u^j d^{n-j}$  for  $0 \leq j \leq n$  and for T. As such, formula (7) above can be represented as:

$$P^{th} = \sum_{j=0}^n A_j \max[SB_j - K, 0] \quad (12)$$

While formulas (7) and (12) operate to determine a theoretical fair market price for a European call options, it should be noted

that this the formula for computing the theoretical fair market price of a European put option can be expressed as:

$$P^{th} = \sum_{j=0}^n p^j q^{n-j} b_j \max[K - Su^j d^{n-j}, 0]$$

(13)

5 which in turn reduces to:

$$P^{th} = \sum_{j=0}^n A_j \max[K - SB_j, 0] \quad (14)$$

With the arrangements of both formulas (12) and (14), each entry 616 in table 610 will contain a pair 612 represented by  $(A(\sigma, T), B(\sigma, T))$ . The lookup table 610 of pairs 612 can be computed at  
 10 scheduled intervals, for example nightly or on a more frequent basis, such as once every two seconds (helping increase the accuracy of the model in the event of changes in parameters such as the risk-free interest rate) throughout the trading day or other time intervals. The architecture can also be configured to update the  
 15 lookup table entries in response to a trigger, such as a change in the risk-free interest rate during the trading day.

Thus, with reference to Figures 6(a) and (b), given an option message 260, a lookup unit 620 preferably operates to parse the message 260 to ascertain the identity of the underlying financial  
 20 instrument (from field 262 of the option message) and the time to maturity parameter therein (from field 268). A lookup table 602 accessible to unit 620 stores volatility values for various financial instruments. As shown by Figure 6(b), lookup table 602 is preferably indexed by financial instrument such that the lookup unit  
 25 620 can retrieve the volatility measure 614 pertinent to option message 260 via a financial instrument index 604. The table 602 of volatility values can be computed on either a continuous basis, at scheduled intervals (e.g., nightly, hourly, etc.), and/or on a triggered basis. The volatility values in table 602 could be  
 30 determined in a number of ways - they can be the underlying financial instruments' historical volatilities computed using historical volatility values for the underlying financial instrument or they can be previously-determined implied volatilities for the underlying financial instrument.

Upon retrieval of the volatility value from lookup table 602, lookup unit 620 then preferably accesses the lookup table 610 using volatility index 606 and time to maturity index 608 taken from message 260. Indices 606 and 608 identify an entry 616 within table 610 that contains an (A,B) pair 612. Lookup unit 620 then retrieves the identified (A,B) pair 612 and passes those values to a combinatorial logic stage 622 to compute the fair market option price 624 for the option message 260.

Figure 6(c) illustrates an exemplary combinatorial logic stage 622 for realizing formulas (12) and (14) above. Preferably, combinatorial logic stage 622 utilizes a plurality  $n+1$  of parallel computational pipelines  $650_0, 650_1, \dots, 650_n$ . Each pipeline  $650_j$  is fed by the following input parameters:  $A_j, B_j, S, K$  and the call/put flag C/P. The C/P flag controls whether formula (12) or formula (14) applies to the computation via multiplexers 650 and 652. To create the  $B_j S$  term of formulas (12) and (14), multiplier 630 operates to multiply  $B_j$  by  $S$ . To create the  $B_j S - K$  term of formula (12), adder 632 preferably adds an inverted  $K$  value to the  $B_j S$  term. To create the  $K - B_j S$  term of formula (14), adder 632 preferably adds the  $K$  value to an inverted value of the  $B_j S$  term. Multiplexers 650 and 652 control which terms reach the adder 632 based on the value of the C/P flag. Next, comparator 634 and multiplexer 636 operate to perform the maximum operation between the terms  $B_j S - K$  and 0 for call options and the terms  $K - B_j S$  and 0 for put options, wherein the maximum of these two values is fed to multiplier 638, whereat the maximum value passed by multiplexer 636 is multiplied by  $A_j$ . Thus, each pipeline  $650_j$  preferably outputs the value of  $A_j \max[B_j S - K, 0]$  for call options or the value of  $A_j \max[K - B_j S, 0]$  for put options to adder 640, to thereby compute the theoretical fair market option price 624.

With the embodiment of Figure 6(c), the computation of fair market option price 624 preferably utilizes  $2(n+1)$  32-bit floating point multiplications,  $(n+1)$  32-bit floating point subtractions,  $(n+1)$  maximum computations, and  $n$  32-bit floating point additions; a total of approximately  $5n$  floating-point operations. For a software embodiment, if one assumes that each operation requires 5 cycles of

a 2 GHz processor, then each operation requires 2.5 ns. If one further assumes that 256 time steps are sufficient to approximate the lognormal distribution, then a fair market option price 624 can be computed in 3.2  $\mu$ s. Further assuming no overhead between fair  
5 market price calculations, then a single 2 GHz processor should be able to make 312,500 fair market option price calculations per second.

Turning to the amount of memory required for the tables 602 and 610, it is preferred that each element of the arrays A and B be  
10 represented as a 32-bit single-precision floating point number. Continuing with the assumption that  $n=255$  is sufficient, then each entry 616 in table 610 will store 256 32-bit values for a total of 8192 bits or 1024 bytes.

If one also assumes that three significant digits are  
15 sufficient to accurately represent the volatility values, this will result in the volatility of any underlying financial instrument assuming one of 1000 possible values. If one further assumes that there are 10 possible maturity dates for all currently traded options (although other numbers of possible maturity dates could be  
20 used), then table 610 will contain 1000x10 entries, which results in a total table size of 102 Mbytes. It is expected that the size of the volatility table 602 will be on the order of 10 kbytes. These table sizes suggest that the lookup tables 602 and 610 can readily be stored in RAM and preferably in a high speed memory, either  
25 processor memory or off-chip memory (for any embodiments employing a hardware co-processor). However, it should also be noted that the lookup table(s) can also be stored in on-chip memory (such as SDRAM on an FPGA chip) if the available memory resources on the chip are sufficient.

30 As for the regularity with which the lookup tables are computed, it is believed that any of a variety of schemes could be employed depending upon the needs and desires of a practitioner of this embodiment of the invention. For example, in some situations, it may be sufficient to re-compute the tables prior to the beginning  
35 of each trading day, particularly if one assumes that the time to maturity of the option can be measured in days and the risk-free

interest rate will not change significantly throughout the trading day.

If a more frequent updating of the lookup tables is desired, then it is helpful to estimate the rate at which the tables can be recomputed. In order to compute the entries in table 610, one must first compute the parameters  $u$ ,  $d$ ,  $p$  and  $q$ . It will be assumed that the value of  $1/n$  will be constant. It will further be assumed that the volatility  $\sigma$  and the risk-free interest rate  $r$  can vary.

Computation of  $u$  requires a floating point square root operation, multiplication operation and exponential operation. Once  $u$  is known, the value of  $d$  can be found by a single floating point divide operation. Computation of  $p$  requires a floating point divide operation, a multiplication operation and two subtraction operations. Once  $p$  is known, the value of  $q$  can be found via a single floating point subtraction operation. The inventors herein estimate that the computation of the  $u$ ,  $d$ ,  $p$  and  $q$  terms can be performed in approximately 200 cycles. If one maximizes the use of intermediate results by using the fully partitioned binomial tree technique exemplified in Figure 10, then a forward pass through the binomial tree requires  $n \lceil \log_2 n \rceil$  multiplication operations. Since two passes are required (once for the array of price factors and once for the array of probabilities), then  $2n \lceil \log_2 n \rceil$  multiplication operations would be needed. Using the same assumptions mentioned above ( $n = 256$  with 2.5 ns required for each multiplication operation), the computation of a lookup table 610 with  $1000 \times 100$  entries would require approximately 1.024 seconds. Accordingly, if a second processor with access to the same system memory is available for continually updating the tables, the inventors herein believe that the table 610 can be re-computed approximately every 2 seconds.

An alternative to the exemplary embodiment of Figures 6(a)-(c) for computing an option's theoretical fair market price is to directly compute the binomial tree used in the fair market price computation (rather than using a lookup table for the terms related to  $u$ ,  $d$ ,  $p$  and  $q$ ).

One approach to such direct computation can utilize the architecture of Figure 5(a). With such a direct computation embodiment, the volatility value used in the theoretical fair market price computation will need to be passed to the OPM computational unit 300, either from a lookup table of volatility values as shown in connection with Figure 6(b) or otherwise.

Figure 7 illustrates an exemplary architecture for another approach to the direct computation of a binomial tree for pricing an option. Once again, this example will be described in connection with a European call option, although it should be understood that other option types can readily be employed in the practice of this embodiment of the invention. As explained herein, the architecture of Figure 7 seeks to take advantage of parallelism to accelerate the fair market price computation. With the architecture of Figure 7, option message 260 is fed into two parallel computational modules 702 and 704. Computational modules 702 and 704 are configured to receive the option characteristics data from message 260 and other pricing parameters, namely  $\sigma$ ,  $r$  and  $n$  so that it can perform the computations necessary for formulas (1), (2), and (8)-(11) to enable an ultimate computation in accordance with formula (7). The volatility input value can be determined from a lookup table such as that shown in Figure 6(b) or otherwise. Computational module 702 is configured to compute the  $p^j q^{n-j} b_j$  terms found in formula (7). Computational module 704 is configured to compute the  $u^j d^{n-j} S$  terms found in formula (7). Downstream from modules 702 and 704 is a combinatorial logic stage 706 that operates on the outputs of modules 702 and 704 as well as the option message 260 to compute the fair market option price 624.

Figure 8(a) illustrates an exemplary binomial tree representation 800 of the  $u^j d^{n-j} S$  terms found in formula (7), wherein  $n=4$ . Each edge 804 in the tree 800 represents a multiplication operation, with each node 802 representing one of the  $u^j d^{n-j} S$  terms. Due to the simplifying assumption that the up and down factors  $u$  and  $d$  for the financial instrument price remain constant over the life of the option, the binomial tree 800 creates a lattice where each node 804 has two "parent" edges 804. One can observe that the price

at each node 802 is independent of the path traveled to reach that node. One can also observe that the terminal prices are the only prices included in the final weighted sum in formula (7).

Therefore, one only needs to compute one path to each terminal price.

Figure 8(b) illustrates one technique for computing a single path to each terminal price, wherein this construction is particularly useful when only  $n+1$  parallel multipliers are available. As shown in Figure 8(d), the tree of Figure 8(a) can be realized in computational module 704 using a plurality of pipelined stages 814. A first stage 814 is fed by the financial instrument price  $S$  and comprises two parallel multiplication units 810 and 812 to compute the terms  $uS$  and  $dS$  (the multipliers 810 and 812 of module 704 also receive the computed  $u$  and  $d$  values respectively for performance of the multiplication operations shown in Figure 8(d)). The second stage 814 comprises three parallel multiplication units (a  $*u$  multiplication unit 810 and two  $*d$  multiplication units 812) to compute the  $u^2S$ ,  $udS$ , and  $d^2S$  terms. The third stage 814 comprises four parallel multiplication units (a  $*u$  multiplication unit 810 and three  $*d$  multiplication units 812) to compute the  $u^3S$ ,  $u^2dS$ ,  $ud^2S$ , and  $d^3S$  terms. The final fourth stage 814 comprises five parallel multiplication units (a  $*u$  multiplication unit 810 and four  $*d$  multiplication units 812) to compute the terminal terms  $u^4S$ ,  $u^3dS$ ,  $u^2d^2S$ ,  $ud^3S$ , and  $d^4S$  terms. Thus, the computational module 704 of Figure 8(d) uses a process of "spawning" new iterative multiplication operations using intermediate results until  $n+1$  parallel multipliers in the final stage 814 compute the terminal financial instrument prices in the  $n^{\text{th}}$  cycle.

Figures 8(c) and 8(e) illustrate an exemplary alternate technique for computing a single path to each terminal price, wherein the terminal price path is different than the path of Figures 8(b) and (d) (essentially the roles of multiplication units 810 and 812 have been reversed). Once again, as with the example of Figures 8(b) and (d), the example of Figures 8(c) and (e) uses a value of  $n$  equal to 4.

While the examples of Figures 8(a)-(e) have used a value of  $n=4$ , it should be understood that a practitioner of this embodiment of the invention can readily use other values of  $n$ . It should be noted that a practitioner of the present invention can select a value for  $n$  based on the speed and accuracy needs of a trader. Larger values of  $n$  will produce a higher degree of accuracy with respect to theoretical fair market price computations, but will increase the time needed to compute the theoretical fair market prices. As such, the value of  $n$  can be set and adjusted as needed to tailor the options pricing techniques described herein to the needs of a given application.

If  $n+1$  parallel multipliers are not available, and/or one wishes to balance the computational load of module 704 across superscalar processors and hardware co-processors, one can partition the binomial tree. One can partition any subtree of  $m$  steps at step  $\lceil m/2 \rceil$ . Figure 9(a) depicts an example of partitioning a binomial tree 900 at step 5. It can be noted that the portion of the binomial tree 900 to the left of partition 902 requires only two paths to compute the terminal prices for the left half of tree 900. These left half computations can be performed sequentially or by parallel processors/threads. Figure 9(b) illustrates an example of a module 910 for computing the terminal prices of the left subtree using parallel multipliers. As shown in Figure 9(b), five pipelined stages 912 of two parallel multiplication units 810 and 812 can be used to compute the terminal  $u^5S$  and  $d^5S$  terms 914 and 916.

At time step 5, the terminal prices 914 and 916 from the left half of tree 900 can be fed to another module, preferably a hardware co-processor. If the hardware co-processor has 10 parallel multipliers available, then the terminal prices  $u^9S$ ,  $u^8dS$ ,  $u^7d^2S$ ,  $u^6d^3S$ ,  $u^5d^4S$ ,  $u^4d^5S$ ,  $u^3d^6S$ ,  $u^2d^7S$ ,  $ud^8S$  and  $d^5S$  can be computed in four time steps. Alternatively, two hardware co-processors with 5 parallel multipliers can be employed to compute those terms in parallel, as shown by modules 920 and 930 in Figures 9(c) and 9(d). In module 920 of Figure 9(c), four pipelined stages 922 compute the terminal prices  $u^9S$ ,  $u^8dS$ ,  $u^7d^2S$ ,  $u^6d^3S$  and  $u^5d^4S$  from the input  $u^5S$  price 914. In module 930 of Figure 9(d), four pipelined stages 932

compute the terminal prices  $u^4d^5S$ ,  $u^3d^6S$ ,  $u^2d^7S$ ,  $ud^8S$  and  $d^9S$  from the input  $d^5S$  price 916. Thus, with the example, computational module 704 can be readily formed from modules 910, 920 and 930. As another alternative, a hardware co-processor with 5 parallel multipliers can compute the terminal prices in 8 steps by sequentially computing each subtree to the right of partition 902 shown in Figure 9(a). It should further be noted that superscalar microprocessor(s) can be used to perform other operations that may be needed for options pricing or for other analytical functions in parallel using available resources on the superscalar microprocessor(s).

It should also be noted that the partitioning technique can be applied recursively to better map the computational load to the particular processing platform used in the practice of this embodiment of the invention. Figure 10(a) depicts an exemplary 11 step binomial tree 1000 where double partitioning has been employed; with one partition 1002 at time step 6 and another partition 1004 at time step 9. In one embodiment, at time step 9, the four intermediate results produced at partition 1004 can be fed to a hardware co-processor, where the hardware co-processor is configured to compute one or more subtrees to the right of partition 1004 in parallel. In another exemplary embodiment drawing from the example of Figure 10(a), module 704 can be realized as shown in Figure 10(b). Partitioned computational unit 1012 outputs the prices  $u^6S$  and  $d^6S$  using two computational paths. The  $u^6S$  term can then be fed to partitioned computational unit 1014, to thereby compute the terms  $u^9S$  and  $u^6d^3S$  using two computational paths. In parallel with partitioned computational unit 1014, partitioned computational unit 1016 can operate on the  $d^6S$  term from unit 1012 to thereby compute the terms  $u^3d^6S$  and  $d^9S$  using two computational paths. Thereafter, the terms  $u^9S$ ,  $u^6d^3S$ ,  $u^3d^6S$  and  $d^9S$  can be fed to four parallel partitioned computational units 1018, 1020, 1022 and 1024 to compute the terminal prices shown in tree 1000 of Figure 10(a).

It should be noted that the same techniques described in connection with Figures 8(a)-10(b) can be used to compute the  $p^j q^{n-j} b_j$  terms via module 702. Figure 11 depicts an exemplary embodiment of module 702 wherein no partitioning has been used and wherein the

binomial tree is of depth  $n=4$ . Module 702 comprises  $n$  pipelined stages 1116 to compute the  $p^j q^{n-j}$  terms via multiplication units 1102 and 1104 (multiplication units 1102 being  $*q$  multipliers and multiplication units 1104 being  $*p$  multipliers). An input value 1118 of 1 is fed into the first pipeline stage 1116. The final stage 1116 comprises  $n+1$  parallel multiplication units 1106, 1108, 1110, 1112, and 1114 that perform the  $*b_j$  multiplication operations. It should be noted that more complex implementations of modules 702 can be deployed in accordance with the teachings set forth herein in connection with Figures 8(a) through 10(b).

It should also be noted that the combinatorial logic stage 706 of architecture 700 of Figure 7 can be constructed as shown in Figure 6(c) for the lookup table embodiment, although the multipliers 630 would not be needed as the operations performed by multipliers 630 would be performed within module 704.

The inventors herein note that the options pricing techniques described in connection with Figures 1(a)-11 can be implemented on any of a number of platforms, in software, hardware, or a combination of hardware and software. As such, as previously explained, suitable platforms for deploying the options pricing techniques described herein can include reconfigurable logic devices (e.g., FPGAs), superscalar processors, multicore processors, ASICs, GPUs, PPUs, chip multi-processors, and GPPs.

A preferred platform for the deployment of the options pricing techniques described herein is the market data platform described in the above-referenced and incorporated provisional United States patent application 60/814,796, an example of which is shown in Figure 13. The market data platform 1300 shown in Figure 13 consolidates the functional units 1202 shown in Figure 12 into much fewer physical devices and also offloads much of the data processing performed by the GPPs of the functional units 1202 to hardware, preferably reconfigurable logic.

For example, with the architecture of Figure 13, the feed compressor 1302 can be deployed in an appliance such as system 1400, described hereinafter with respect to Figure 14. Feed compressor 1302 is used to compress the content of the financial data stream

1206 arriving from various individual sources. Examples of compression techniques that can be used include the open standard "glib" as well as any proprietary compression technique that may be used by a practitioner of the present invention.

5 Preferably, the feed compressor device 1302 is deployed in a physical location as close to the feed source 1206 as possible, to thereby reduce communication costs and latency. For example, it would be advantageous to deploy the feed compressor device 1302 in a data center of an extranet provider (e.g., Savvis, BT Radians, etc.)  
10 due to the data center's geographic proximity to the source of the financial market data 1206. Because the compression reduces message sizes within the feed stream 1206, it will be advantageous to perform the compression prior to the stream reaching wide area network (WAN) 1320a; thereby reducing communication latency through  
15 the network because of the smaller message sizes.

WAN 1320 preferably comprises an extranet infrastructure or private communication lines for connection, on the inbound side, to the feed handlers deployed in device 1304. On the outbound side, WAN 1320 preferably connects with device 1306, as explained below.  
20 It should be noted that WAN 1320 can comprise a single network or multiple networks 1320a and 1320b segmented by their inbound/outbound role in relation to platform 1300. It is also worth noting that a news feed with real-time news wire reports can also be fed into WAN 1320a for delivery to device 1304.

25 Device 1304 can be deployed in an appliance such as system 1400 described hereinafter with respect to Figure 14. Whereas the conventional GPP-based system architecture shown in Figure 12 deployed the functional units of feed handling/ticker plant, rule-based calculation engines, an alert generation engine, options  
30 pricing, last value cache (LVC) servers supplying snapshot and/or streaming interfaces, historical time-series oriented databases with analytics, and news databases with search capabilities in software on separate GPPs, the architecture of Figure 13 can consolidate these functions, either partially or in total, in firmware resident  
35 on the reconfigurable logic (such as one or more FPGAs) of device 1304.

Feed handlers, LVCs, time series databases, and news databases can be implemented in device 1304 as described in connection with the above-referenced and incorporated 60/814,796 application.

5 The rule-based calculation engines, which are engines amenable to implementation in a firmware pipeline and that allow a user to create his/her own synthetic records whose field values are derived from calculations performed against information obtained from the LVC, information extracted from a stream of update messages generated from the LVC, or from alternate sources, can optionally  
10 employ at least in part the options pricing techniques described herein. It should also be noted that the rule-based calculation engine can be configured to create new synthetic fields that are included in existing records maintained by the LVC. The new values computed by the engine are computed by following a set of rules or  
15 formulas that have been specified for each synthetic field. Thus, for example, a rule-based calculation engine can be configured to compute an option's implied volatility and/or theoretical fair market price using the techniques described in connection with Figures 1(a)-11 and include those values in existing records for the  
20 option maintained by the LVC.

The alert generation functionality within device 1304 can also be deployed in a firmware pipeline. Alert generation engines are similar to a rule-based calculation engine in that they monitor the current state of a financial instrument record (or set of financial  
25 instrument records), and the alert generation engine will trigger an alert when any of a set of specified conditions is met (e.g., when an option's actual price  $P$  is within some pre-specified tolerance of an option's theoretical fair market price). An indication is then delivered via a variety of means to consuming applications or end  
30 users that wish to be notified upon the occurrence of the alert. This embodiment of the present invention provides great value by providing a practitioner with the ability to tightly couple alert generation functionality with other applications that react to the alerts generated thereby and reduce the effective latency of data  
35 processing (e.g., a "black box" trading algorithm that may respond to an alert detected by the alert generation engine).

Options pricing can be implemented with a firmware pipeline in device 1304, via either the rule-based calculation engine functionality described above or via a specific options pricing engine. Thus, an option pricing engine that is part of device 1304 can be configured to perform a number of computations related to received options and their underlying instruments (e.g., the theoretical fair market value of an option or the implied volatility of the underlying instrument based upon the market price of the option as described in connection with Figures 1(a)-11). However, it should be noted that the options pricing engine need not be limited to the techniques described herein with respect to Figures 1(a)-11, as a wide array of computational rules can be used for pricing options, as is known in the art, which can be deployed in firmware for platform 1300. As explained above, most if not all industry-accepted techniques for options pricing are extremely computation intensive which introduces significant latency when the computations are performed in software. However, by implementing option pricing in a firmware pipeline, the market data platform 1300 can significantly accelerate the computation of option pricing, thereby providing an important edge to traders who use the present invention.

Traders at workstations 1204 (or application programs 1350 running on an entity's own trading platform) can then access the streaming financial data processed by device 1304 via a connection to local area network (LAN) 1322. Through this LAN connection, workstations 1204 (and application program 1350) also have access to the data produced by devices 1306, 1308, 1310, 1312, 1314, and 1316. Like devices 1302 and 1304, devices 1306, 1308, 1310, 1312, 1314, and 1316 can also be deployed in an appliance such as system 1400 described hereinafter with respect to Figure 14.

As described in the above-referenced and incorporated 60/814,796 application, device 1306 preferably consolidates the following functionality at least partially into firmware resident on reconfigurable logic: an order book server; an order router; direct market access gateways to exchanges, Electronic Communication Networks (ECNs), and other liquidity pools; trading engines; an

auto-quote server; and a compliance journal. An order book server can employ an options pricing engine or the results from an options pricing engine to enrich the order book data with computed theoretical fair market prices and/or implied volatilities for the options. Also, an order router can be enhanced by employing an options pricing engine or the results from an options pricing engine. For example, the order router may be configured to monitor the status of the order book. In doing so, the order router can identify where liquidity exists in the market to help drive its order routing decisions. When making such order routing decisions, knowledge of an option's implied volatility and/or theoretical fair market price may be helpful. A feed/compliance journal could also employ an options pricing engine or the results from an options pricing engine to correlate market prices with theoretical fair market prices and/or implied volatilities to identify any potential market abnormalities.

A trading engine within device 1306 can also be deployed on reconfigurable logic. An algorithmic trading engine operates to apply a quantitative model to trade orders of a defined quantity to thereby automatically subdivide that trade order into smaller orders whose timing and size are guided by the goals of the quantitative model so as to reduce the impact that the original trade order may have on the current market price. Also, a black box trading engine operates to automatically generate trades by following a mathematical model that specifies relationships or conditional parameters for an instrument or set of instruments. To aid this processing, the black box trading engine is fed with real-time market data. Thus, the black box trading engine within device 1306 can be fed with the computed implied volatility and/or computed fair market price(s) for each option so that the black box trading engine can make a decision on whether to buy or sell a given option. This combination of a black box trading engine with the accelerated options pricing functionality described herein synergistically creates incredible competitive advantages for practitioners of this embodiment of the invention in that the practitioner can conceivably take action on options available on the market before other traders

even realize whether the available option represents a desirable deal or not.

An auto-quote server is similar to a black box trading engine. The auto-quote server operates to automatically generate firm quotes to buy or sell a particular financial instrument at the behest of a "market maker"; wherein a "market maker" is a person or entity which quotes a buy and/or sell price in a financial instrument hoping to make a profit on the "turn" or the bid/offer spread. By employing an options pricing engine or the results from an options pricing engine, an auto-quote server's decisions as to what bid/ask offers to place in the market can be driven at least in part by the computed theoretical fair market prices and/or implied volatilities. For example, it is conceivable that the bid/offer spread may correlate to an instrument's volatility; these and other relationships could be examined by a black-box auto-quote server that employs an options pricing engine or the results therefrom.

As described in the above-referenced and incorporated 60/814,796 application, device 1308 preferably implements an internal matching system/engine in firmware resident on reconfigurable logic, device 1310 preferably implements an order management system (OMS) in firmware resident on reconfigurable logic, device 1312 preferably implements entitlements and reporting functionality, device 1314 preferably implements management and monitoring, and device 1316 preferably implements publishing and contribution server functionality.

Figure 14 depicts an exemplary system 1400 in which the options pricing functionality described herein (and optionally also the other functionality described in connection with Figure 13) can be deployed. The system 1400 can be characterized as a hardware co-processor or hardware appliance that performs high speed processing of streaming data. For e.g., the processing may include financial calculations such as the implied volatility or options price computations mentioned previously. In this system, a reconfigurable logic device 1402 is positioned to receive data that streams off of either or both a disk subsystem defined by disk controller 1406 and data store 1404 (either directly or indirectly by way of system

memory such as RAM 1410) and a network data source/destination 1442 (via network interface 1440). Preferably, data streams into the reconfigurable logic device by way of system bus 1412, although other design architectures are possible (see Figure 16(b)).

5 Preferably, the reconfigurable logic device 1402 is an FPGA, although this need not be the case. For example, the reconfigurable logic device may also take the form of generalized field programmable object array (FPOA) wherein the objects may be processing elements (e.g., arithmetic units) or full processors  
10 (e.g. chip-multi-processors). System bus 1412 can also interconnect the reconfigurable logic device 1402 with the computer system's main processor 1408 as well as the computer system's RAM 1410. The term "bus" as used herein refers to a logical bus which encompasses any physical interconnect for which devices and locations are accessed  
15 by an address. Examples of buses that could be used in the practice of the present invention include, but are not limited to the PCI family of buses (e.g., PCI-X and PCI-Express) and HyperTransport buses. In a preferred embodiment, system bus 1412 may be a PCI-X bus, although this need not be the case.

20 The data store can be any data storage device/system, but is preferably some form of a mass storage medium. For example, the data store 1404 can be a magnetic storage device such as an array of hard-disk drives. However, it should be noted that other types of storage media may also be suitable for use in the practice of the  
25 invention. For example, the data store could also be one or more remote data storage devices that are accessed over a network such as the Internet, a storage area network (SAN), or some local area network (LAN). Another source/destination for data streaming to or from the reconfigurable logic device 1402, is network 1442 by way of  
30 network interface 1440, as described above. In the financial industry, a network data source (e.g., the exchanges themselves, a third party provider, etc.) can provide the financial data stream 1206 described above in connection with Figures 12 and 13.

The computer system defined by main processor 1408 and RAM  
35 1410 is preferably any commodity computer system as would be understood by those having ordinary skill in the art. For example,

the computer system may be an Intel Xeon system or an AMD Opteron system.

The reconfigurable logic device 1402 has firmware modules deployed thereon that define its functionality. In one  
5 instantiation, the firmware socket module 1420 handles the data movement requirements (both command data and target data) into and out of the reconfigurable logic device, thereby providing a consistent application interface to the firmware application module (FAM) chain 1430 that is also deployed on the reconfigurable logic  
10 device. The FAMs 1430i of the FAM chain 1430 are configured to perform specified data processing operations on any data that streams through the chain 1430 from the firmware socket module 1420. Preferred examples of FAMs that can be deployed on reconfigurable logic in accordance with a preferred embodiment of the present  
15 invention are described above in connection with options pricing. For example, the pipeline 250 of Figure 2(b) (including the components therein as embodied by the examples of Figure 3 and 5) can be deployed in a FAM chain 1430 to compute an option's implied volatility. The combinatorial logic stage 622 of Figure 6(a) can  
20 also be deployed in a FAM chain 1430 to compute an option's fair market price if desired by a practitioner of the invention. Furthermore, it should be noted that the lookup unit 620 and one or more of the lookup tables described in connection with Figures 6(a) and (b) could be deployed in the FAM chain 1430 if there are  
25 sufficient resources available on the reconfigurable logic. Similarly, all or portions of the architecture 700 described by Figure 7 can be deployed in a FAM chain 1430 to compute an option's theoretical fair market price.

The specific data processing operation that is performed by a  
30 FAM can be controlled/parameterized by the command data that FAM receives from the firmware socket module 1420. This command data can be FAM-specific, and upon receipt of the command, the FAM will arrange itself to carry out the data processing operation controlled by the received command. For example, a FAM that is configured as  
35 the initial module 252 of pipeline 250 of Figure 2(b) can be parameterized to define the values for  $n$ ,  $r$ ,  $\epsilon$ , and  $\epsilon_0$  and the

initial volatility band boundary values  $\sigma^1_1, \sigma^1_2, \dots \sigma^1_{m+1}$ . In this way, a FAM that is configured to compute a specified data value for a given option can be readily re-arranged to compute a data value for a different option by simply loading new parameters for the different option in that FAM. In this manner, a high throughput can be maintained as multiple options are processed through FAM chain 1430.

Once a FAM has been arranged to perform the data processing operation specified by a received command, that FAM is ready to carry out its specified data processing operation on the data stream that it receives from the firmware socket module. Thus, a FAM can be arranged through an appropriate command to process a specified stream of data in a specified manner. Once the FAM has completed its data processing operation, another command can be sent to that FAM that will cause the FAM to re-arrange itself to alter the nature of the data processing operation performed thereby. In a preferred embodiment of the system, a FAM may process commands in parallel with data streams. For example, new parameters for the next stream of data may be loaded into a parameter cache prior to completion of the current data stream. Not only will the FAM operate at hardware speeds (thereby providing a high throughput of data through the FAM), but the FAMs can also be flexibly reprogrammed to change the parameters of their data processing operations.

The FAM chain 1430 preferably comprises a plurality of firmware application modules (FAMs) 1430a, 1430b, ... that are arranged in a pipelined sequence. As used herein, "FAM pipeline", "FAM pipelined sequence", or "FAM chain" refers to an arrangement of FAMs wherein the output of one FAM is connected to the input of the next FAM in the sequence. This pipelining arrangement allows each FAM to independently operate on any data it receives during a given clock cycle and then pass its output to the next downstream FAM in the sequence during another clock cycle.

A communication path 1432 connects the firmware socket module 1420 with the input of the first one of the pipelined FAMs 1430a. The input of the first FAM 1430a serves as the entry point into the FAM chain 1430. A communication path 1434 connects the output of

the final one of the pipelined FAMs 1430z with the firmware socket module 1420. The output of the final FAM 1430z serves as the exit point from the FAM chain 1430. Both communication path 1432 and communication path 1434 are preferably multi-bit paths.

5        Figure 15 depicts an exemplary framework for the deployment of applications on the system 1400 of Figure 14. The top three layers of Figure 15 represent functionality that is executed in software on the computer system's general-purpose processor 1408. The bottom  
10       two layers represent functionality that is executed in firmware on the reconfigurable logic device 1402.

      The application software layer 1500 corresponds to high level functionality such as the type of functionality wherein one or more users or external programs interact with the application to define which data processing operations are to be performed by the FAMs and  
15       to define what data those data processing operations are to be performed upon.

      The next layer is the module application programming interface (API) layer 1502 which comprises a high level module API 1502a and a low level module API 1502b. The high level module API 1502a can  
20       provide generic services to application level software (for example, managing callbacks). The low level module API 1502b manages the operation of the operating system (OS) level/device driver software 1504. A software library interface 1510 interfaces the high level module API 1502a with the low level module API 1502b. Additional  
25       details about this software library interface can be found in the above-referenced and incorporated patent application 11/339,892.

      The interface between the device driver software 1504 and the firmware socket module 1420 serves as the hardware/software interface 1512 for the system 1400. The details of this interface  
30       1512 are described in greater detail in the above-referenced and incorporated patent application 11/339,892.

      The interface between the firmware socket module 1420 and the FAM chain 1430 is the firmware module interface 1514. The details of this interface are described in greater detail in the above-  
35       referenced and incorporated patent application 11/339,892.

Figure 16(a) depicts a printed circuit board or card 1600 that can be connected to the PCI-X bus 1412 of a commodity computer system for use in a market data platform. In the example of Figure 16(a), the printed circuit board includes an FPGA 1602 (such as a  
5 Xilinx Virtex 4 FPGA) that is in communication with a memory device 1604 and a PCI-X bus connector 1606. A preferred memory device 1604 comprises SRAM and DRAM memory. A preferred PCI-X bus connector 1606 is a standard card edge connector.

Figure 16(b) depicts an alternate configuration for a printed  
10 circuit board/card 1600. In the example of Figure 16(b), a private bus 1608 (such as a PCI-X bus), a network interface controller 1610, and a network connector 1612 are also installed on the printed circuit board 1600. Any commodity network interface technology can be supported, as is understood in the art. In this configuration,  
15 the firmware socket 1420 also serves as a PCI-X to PCI-X bridge to provide the processor 1408 with normal access to the network(s) connected via the private PCI-X bus 1608.

It is worth noting that in either the configuration of Figure 16(a) or 16(b), the firmware socket 1420 can make memory 1604  
20 accessible to the PCI-X bus, which thereby makes memory 1604 available for use by the OS kernel 1504 as the buffers for transfers from the disk controller and/or network interface controller to the FAMS. It is also worth noting that while a single FPGA 1602 is shown on the printed circuit boards of Figures 16(a) and (b), it  
25 should be understood that multiple FPGAs can be supported by either including more than one FPGA on the printed circuit board 1600 or by installing more than one printed circuit board 1600 in the computer system. Figure 17 depicts an example where numerous FAMS in a single pipeline are deployed across multiple FPGAs.

30 As shown in Figures 14-16(b), inbound data (from the kernel 1504 to the card 1600) is moved across the bus 1412 in the computer system to the firmware socket module 1420 and then delivered by the firmware socket module 1420 to the FAM chain 1430. Outbound data (from the card 1600 to the kernel 1504) are delivered from the FAM  
35 chain 1430 to the firmware socket module 1420 and then delivered by the firmware socket module 1420 across the PCI-X bus to the software

application executing on the computer system. As shown in Figure 15, the three interacting interfaces that are used are the firmware module interface 1514, the hardware/software interface 1512, and the software library interface 1510.

Figure 18 depicts an exemplary FAM pipeline for carrying out a variety of data processing tasks on streaming financial information, wherein the options pricing functionality described herein can be deployed in whole or in part within the field calculation engine 1812. The FAM pipeline of Figure 18 takes in a FAST message stream.

FAST (FIX Adapted for Streaming) is known in the art as a message encoding scheme. The incoming FAST stream is received by FAM 1802, which deserializes the stream of FAST messages. The deserialized FAST messages are then provided to FAM 1804, which operates to decode the various fields of the FAST messages as aided by the templates and field maps in memory 1822 and the operator values in memory 1824. Thus, the output of FAM 1804 comprises the data content of the FAST message decomposed into its constituent fields.

This content is then passed to a variety of parallel FAMS 1806, 1808, and 1810. FAM 1806 performs a administrative record filter on the data it receives. The administrative record filter preferably operates to pass through message types that are not processed by any of the other FAM modules of the pipeline. FAM 1808 serves as a Top 10 lists engine, as described in the above-referenced and incorporated 60/814,796 application. FAM 1810 serves as a message

query filter. Message query filters allow for certain messages to be excluded from the message flow. Such filters are preferably parameterized in FAM 1810 such that filtering criteria based on the field values contained within each message can be flexibly defined and loaded onto the FPGA. Examples of filtering criteria that can be used to filter messages include a particular type of instrument (e.g., common stock, warrant, bond, option, commodity, future, etc.), membership within a prescribed set of financial instruments (e.g., an index or "exchange traded fund" (ETF)), message type, etc. In the embodiment herein wherein the field calculation engine 1812 comprises an options pricing engine, the message query filter 1810

is preferably configured to forward option messages to the field calculation engine for processing thereby.

Thus, as noted, the output of FAM 1810 (which comprises a stream of option messages) is then passed to FAM 1812, which is configured as a rule-based calculation engine to perform options pricing. FAM 1812 also preferably receives data from a real time field value cache 1826 to obtain LVC data, as does the top 10 list FAM 1808. Cache 1826 is preferably embodied by storage 1332. The output from the rule-based calculation engine FAM 1812 is then passed to parallel FAMs 1814, 1816, and 1818. FAM 1814 can serve as a message multiplexer, and receives messages from the outputs of FAMs 1806, 1808 and 1812. FAM 1820 receives the messages multiplexed by FAM 1814, and serves to encode those messages to a desired format. FAM 1816 serves as an alert generation engine, whose function is explained above, and whose output exits the pipeline. FAM 1818 serves as a value cache update engine to ensuring that cache 1826 stays current.

Figure 19 depicts another exemplary FAM pipeline for carrying out multiple data processing tasks. FAM 1902 takes in a stream of fixed format messages and parses those messages into their constituent data fields. The output of FAM 1902 can be provided to FAM 1904 and FAM 1918. FAM 1918 serves as a message synchronization buffer. Thus, as the fields of the original parsed message are passed directly from FAM 1902 to FAM 1918, FAM 1918 will buffer those data fields while the upper path of Figure 19 (defined by FAMs 1904, 1906, 1908, 1910, 1912, and 1914) process select fields of the parsed message. Thus, upon completion of the processing performed by the FAMs of the upper path, the message formatting FAM 1916, can generate a new message for output from the pipeline using the fields as processed by the upper path for that parsed message as well as the fields buffered in FAM 1918. The message formatter 1916 can then append the fields processed by the upper path FAMs to the fields buffered in FAM 1918 for that message, replace select fields buffered in FAM 1918 for that message with fields processed by the upper path FAMs, or some combination of this appending and replacing.

FAM 1904 operates to map the known symbol for a financial instrument (or set of financial instruments) as defined in the parsed message to a symbology that is internal to the platform (e.g., mapping the symbol for IBM stock to an internal symbol "12345"). FAM 1906 receives the output from FAM 1904 and serves to update the LVC cache via memory 1924. The output of FAM 1906 is then provided in parallel to FAMs 1908, 1910, 1912, and 1914.

FAM 1908 operates as a Top 10 list generator, as described above. FAM 1910 operates as a Minute Bar generator, as described in the above-referenced and incorporated 60/814,796 application. FAM 1912 operates as an interest/entitlement filter, as described in the above-referenced and incorporated 60/814,796 application, and FAM 1914 operates as a programmatic calculation engine, as described above. In this embodiment, the programmatic calculation engine 1914 can employ options pricing as described in connection with any of the embodiments of Figures 1(a)-11. The outputs from FAMs 1908, 1910, 1912 and 1914 are then provided to a message formatter FAM 1916, which operates to construct a fixed format message of a desired format from the outputs of FAMs 1908, 1910, 1912, 1914 and 1918.

In performing these tasks, FAM 1904 is aided by memory 1920 that stores templates and field maps, as well as memory 1922 that stores a symbol index. FAM 1906 is also aided by memory 1920 as well as memory 1924 which serves as an LVC cache. Memory 1920 is also accessed by FAM 1908, while memory 1924 is also accessed by FAM 1914. FAM 1912 accesses interest entitlement memory 1926, as loaded from storage 1334 during initialization of the board 1600.

While these figures illustrate several embodiments of FAM pipelines that can be implemented to process real time financial data streams, it should be noted that numerous other FAM pipelines could be readily devised and developed by persons having ordinary skill in the art following the teachings herein.

Further still it should be noted that for redundancy purposes and/or scaling purposes, redundant appliances 1304, 1306, 1308, 1310, 1312, 1314 and 1316 can be deployed in a given market data platform 1300.

Furthermore, it should also be noted that a practitioner of the present invention may choose to deploy less than all of the functionality described herein in reconfigurable logic. For example, device 1304 may be arranged to perform only options pricing  
5 in reconfigurable logic, or some other subset of the functions listed in Figure 13. If a user later wanted to add additional functionality to device 134, it can do so by simply re-configuring the reconfigurable logic of system 1400 to add any desired new  
10 data processing functionality that can be considered to belong to the same category of data processing operations. That is, devices 1312 and 1314 can be categorized as management operations. Device 1304 can be categorized as providing feed handling/processing for data access, value-added services, and historic services. Devices  
15 1306, 1308 and 1310 can be categorized as direct market access trading systems. As improvements to reconfigurable logic continues over time such that more resources become available thereon (e.g., more available memory on FPGAs), the inventors envision that further consolidation of financial data processing functionality can be  
20 achieved by combining data processing operations of like categories, as indicated by the dashed boxes, thereby further reducing the number of appliances 1400 needed to implement platform 1300. Further still, in the event of such resource improvements over time for FPGAs, it can be foreseen that even further consolidation occur,  
25 including consolidation of all functionality shown in Figure 13 on a single system 1400.

Thus, a platform 1300 developed in the practice of the invention can be designed to improve data processing speeds for financial market information, all while reducing the number of  
30 appliances needed for platform 1300 (relative to conventional GPP-based systems) as well as the space consumed by such a platform. With a platform 1300, a user such as a trader at a work station 1204 (or even a customer-supplied application software program 1350 that accesses the platform via an application programming interface  
35 (API), can obtain a variety of information on the financial markets with less latency than would be expected from a conventional system.

This improvement in latency can translate into tremendous value for practitioners of the invention.

While the present invention has been described above in relation to its preferred embodiment, various modifications may be made thereto that still fall within the invention's scope.

For example, it should be noted that the OPM computational units 300 of Figure 3 can utilize a look-up table approach to computing the theoretical fair market option price for a given input  $\sigma_i^q$ , as shown in Figures 20(a) and (b). Such an OPM computational unit 300' preferably would include a lookup unit 2020 in communication with a combinatorial logic stage 622 like that described in connection with Figure 6(c). The lookup unit 2020 preferably employs a lookup table 610 such as that described in connection with Figure 6(b). Lookup unit 2020 then operates to (1) receive an option message 260 and an input volatility value  $\sigma_i^q$ , (2) parse the option message 260 to identify an index value 2008 for indexing the lookup table 610 (preferably the value of the option's time to maturity characteristic), (3) retrieve the precomputed term(s) 612 (e.g., the (A,B) pair) located in the table entry 616 defined by the time to maturity index 2008 and the volatility index 2006. The retrieved precomputed term(s) 612 can then be passed to the combinatorial logic stage 622, which is configured to compute the option's theoretical fair market price.

As another example, it should be noted that while the embodiment of Figure 2(b) depicts a plurality of pipelined computational modules 252 wherein each computational module 252 is configured to perform a different iteration of the algorithm shown in Figure 2(a), the system can also be configured such that the same computational module 252 performs the computations for all iterations, in which case the output from computational module 252 will be fed back to itself for subsequent iterations. While this configuration will likely degrade the system's throughput capabilities, such a design may be desirable in view of the amount of processing resources available on a processing platform to reduce the system's required amount of computational resources. In a like manner, a hybrid approach can also be taken wherein a plurality of

computational modules 252 are arranged in a pipeline 250, but where a subset of the computational modules 252 (possibly only one thereof) is configured to perform the computations for multiple iterations of the algorithm shown in Figure 2(a). In such a configuration, a balance between throughput and available computational resources can be sought if the amount of available computational resources poses a challenge.

Similarly, it should also be noted that while the embodiment of Figure 3 depicts a plurality of parallel OPM computational units 300 operating within a computational module 252, the computational module 252 can also be configured such that is configured to perform a different iteration of the algorithm shown in Figure 2(a), the system can also be configured with only one OPM computational unit 300, wherein the single computational unit 300 will sequentially compute the theoretical fair option price for each of the different volatility values for a given iteration. Once again, while this design will likely sacrifice throughput, such a design may be desirable in view of the amount of computational resources available on a given processing platform. Also as mentioned above, a hybrid approach can be taken wherein a plurality of parallel OPM computational units 300 are deployed in the computational module 252, but the number of OPM computational units 300 is not sufficient to separately compute the theoretical fair market prices for each volatility value of a given iteration such that at least one of the OPM computational units 300 operates sequentially on a plurality of the different volatility values for that iteration. Preferably, the workload of volatility values for a given iteration would be evenly distributed across the plurality of sequentially operating OPM computational units 300. It should also be noted that in such a design, the combinatorial logic stage 304 would be configured with buffer space in which to store the outputs from the difference computational units 302 while the OPM computational units 300 are sequentially operating on the next set of volatility values for the given iteration.

It should also be noted that, with the preferred embodiment, the step of identifying the volatility band within which the implied

volatility resides comprises defining the band from an upper and lower volatility value for which the computed theoretical fair market prices surround the option's actual purchase price. Another option for identifying the band within which the implied volatility resides is to determine the volatility value corresponding to the computed theoretical fair market price that is closest in value to the option's actual purchase price. The identified band could then be defined by adding some additional width around the determined volatility value in both directions.

Moreover, while the preferred embodiment describes that the same pipeline can be used to process both call and put options as well as both American and European options (wherein appropriate flags within option messages are used to control whether the pipeline treats a given option message as a call/put and/or American/European option), it should be noted that the system can also be configured to maintain multiple separate pipelines, wherein each pipeline is configured to a given type of option (e.g., a two pipeline system wherein one pipeline is for American options and one pipeline is for European options; a two pipeline system wherein one pipeline is for call options and one pipeline is for put options; and a four pipeline system wherein one pipeline is for American call options, one pipeline is for American put options, one pipeline is for European call options, and one pipeline is for European put options). In such an embodiment, option messages can be directed toward an appropriate pipeline based on its call/put flags and/or American/European flags.

These and other modifications to the invention will be recognizable upon review of the teachings herein. As such, the full scope of the present invention is to be defined solely by the appended claims and their legal equivalents.

WHAT IS CLAIMED IS:

1. A device for processing financial market data, the device comprising:
  - 5 a reconfigurable logic device that is configured to (1) receive a data stream comprising financial market data and (2) perform an options pricing operation on at least a portion of the received data stream.
- 10 2. The device as defined in claim 1 wherein the options pricing operation comprises a computation of an implied volatility for an option on an underlying financial instrument.
- 15 3. The device as defined in any of the preceding claims 1-2 wherein the options pricing operation further comprises a computation of at least one theoretical fair market price for the option.
- 20 4. The device as defined in claim 3 wherein the reconfigurable logic device is further configured to compute the at least one theoretical fair market price by applying an option pricing model to at least a portion of the received data stream.
- 25 5. The device as defined in claim 4 wherein the option pricing model comprises an option pricing model based on a Cox-Ross-Rubinstein (CRR) pricing model.
- 30 6. The device as defined in any of claims 1-5 wherein the reconfigurable logic device is configured with a firmware application module pipeline, the firmware application module pipeline being configured to perform the options pricing operation.
- 35 7. The device as defined in claim 2 wherein the implied volatility computation comprises an iterative computation of the implied volatility.

8. The device as defined in any of the preceding claims 1-7 wherein the options pricing operation comprises a computation of a theoretical fair market price for an option on an underlying financial instrument.

5

9. The device as defined in any of the preceding claims 3-5 or 8 wherein the reconfigurable logic device is configured to perform the options pricing operation using a lookup table indexed by a volatility for the option, the lookup table being configured to store a plurality of precomputed parameters for a plurality of different volatility values, the precomputed parameters for use by the reconfigurable logic device in the computation of the theoretical fair market price.

10

10. The device as defined in claim 9 wherein the lookup table is external to the reconfigurable logic device.

15

11. The device as defined in any of the preceding claims 3-5 or 8-10 wherein the reconfigurable logic device is further configured to directly compute a plurality of constituent terms used in the theoretical fair market price computation via a plurality of pipelined stages, wherein a plurality of the pipelined stages comprise a plurality of parallel computational units.

20

12. The device as defined in any of the preceding claims 3-5 or 8-11 wherein the reconfigurable logic device is further configured to perform computation of the theoretical fair market price based on a binomial option pricing model.

25

13. The device as defined in any of the preceding claims 1-12 wherein the reconfigurable logic device comprises a field programmable gate array (FPGA).

30

14. A method for processing financial market data, the method comprising:

35

streaming financial market data through a reconfigurable logic device; and

performing an options pricing operation with the reconfigurable logic device on at least a portion of the financial market data as the financial market data streams therethrough.

15. The method as defined in claim 14 wherein the performing step comprises computing an implied volatility for an option on an underlying financial instrument.

16. The method as defined in any of the preceding claims 14-15 wherein the performing step further comprises computing at least one theoretical fair market price for the option.

17. The method as defined in claim 16 wherein the theoretical price computing step comprises computing the at least one theoretical fair market price by applying an option pricing model to at least a portion of the streaming financial market data.

18. The method as defined in claim 17 wherein the option pricing model comprises an option pricing model based on a Cox-Ross-Rubinstein (CRR) pricing model.

19. The method as defined in claim 15 wherein the computing step comprises computing the implied volatility with a firmware application module pipeline.

20. The method as defined in any of the preceding claims 15 or 19 further comprising:

feeding the computed implied volatility to a black box trading algorithm; and

automatically executing a trade on the option via the black box trading algorithm based at least in part upon the computed implied volatility.

21. The method as defined in any of the preceding claims 15 or 19-20 further comprising:

automatically generating an option quote with an auto-quote server based at least in part upon the computed implied volatility.

5

22. The method as defined in any of the preceding claims 14-21 wherein the performing step comprises computing a theoretical fair market price for an option on an underlying financial instrument.

10 23. The method as defined in any of the preceding claims 16-18 or 22 further comprising:

storing a plurality of precomputed parameters for a plurality of different volatility values in a lookup table indexed by a volatility index; and

15 retrieving a precomputed parameter for use in the computing step from the lookup table using a volatility value for the option as the volatility index into the lookup table; and

wherein the computing step comprises computing the theoretical fair market price based at least in part upon the retrieved precomputed parameter.

20

24. The method as defined in any of the preceding claims 16-18 or 22-23 further comprising:

25 feeding the computed theoretical fair purchase price to a black box trading algorithm; and

automatically executing a trade on the option via the black box trading algorithm based at least in part upon the computed theoretical fair purchase price.

30 25. The method as defined in any of the preceding claims 16-18 or 22-24 further comprising:

automatically generating an option quote with an auto-quote server based at least in part upon the computed theoretical fair purchase price.

35

26. The method as defined in any of the preceding claims 16-18 or 22-25 wherein the computing step comprises directly computing a plurality of constituent terms used in the theoretical fair market price computation via a plurality of pipelined stages, wherein a  
5 plurality of the pipelined stages comprise a plurality of parallel computational units.

27. A device for computing an implied volatility for an option on an underlying financial instrument, the device comprising:

10 a plurality of computational modules arranged in a pipeline, each computational module being configured to (1) receive a plurality of option characteristic values for the option, (2) compute a plurality of theoretical fair market prices for the option from the option characteristic values and a plurality of different  
15 volatility values, (3) determine whether a condition has been satisfied, and (4) in response to a determination that the condition has been satisfied, determine the implied volatility for the option based at least in part on the computed theoretical fair market prices; and

20 wherein each computational module in the pipeline is configured to successively operate on a progressively narrower band of volatility values for the option.

28. The device as defined in claim 27 wherein each computational  
25 module is configured to compute the plurality of theoretical fair market prices in parallel.

29. The device as defined in any of the preceding claims 27-28 wherein the pipeline is configured to process a data stream  
30 comprising a plurality of option messages, each option message corresponding to a different option and comprising a plurality of option characteristic values therefor, and wherein successive computational modules in the pipeline are configured to operate sequentially on the same option message such that each computational  
35 module in the pipeline is configured to operate in parallel on

different option messages to thereby determine the implied volatilities for the different option messages.

30. The device as defined in any of the preceding claims 27-29 wherein the condition comprises a theoretical fair market price convergence property.

31. The device as defined in claim 30 wherein each computational module is further configured to (1) determine whether the theoretical fair market price convergence property condition is satisfied by comparing each theoretical fair market price computed thereby with an actual purchase price for the option and determining that the theoretical fair market price convergence property condition is satisfied if one of the theoretical fair market prices computed thereby differs from the actual purchase price by less than the theoretical fair market price convergence property, and (2) in response to a determination that the theoretical fair market price convergence property condition has been satisfied, determine the implied volatility for the option as the volatility value for which the theoretical fair market price computed therefrom differs from the actual purchase price by less than the theoretical fair market price convergence property.

32. The device as defined in any of the preceding claims 27-31 wherein the condition comprises a volatility convergence property.

33. The device as defined in claim 32 wherein each computational module is further configured to identify a band of volatility values within which the implied volatility resides based at least in part upon the plurality of theoretical fair market prices computed by that computational module, wherein the identified band has a width that is less than a full range of the volatility values for which that computational module computed the plurality of theoretical fair market prices, wherein each computational module is further configured to (1) determine whether the volatility convergence property condition is satisfied by comparing the width of the

identified band with the volatility convergence property, and (2) in response to a determination that the volatility convergence property condition has been satisfied, determine the implied volatility for the option as a value within the identified band.

5

34. The device as defined in any of the preceding claims 27-33 wherein each computational module is further configured to determine whether at least one of at least two conditions has been satisfied, and, in response to a determination that at least one of the  
10 conditions has been satisfied, determine the implied volatility for the option based at least in part on the computed theoretical fair market prices, wherein the at least two conditions comprise a condition corresponding to a theoretical fair market price convergence property and a condition corresponding to a volatility  
15 convergence property.

35. The device as defined in any of the preceding claims 27-34 wherein each computational module is further configured to identify a band of volatility values within which the implied volatility  
20 resides based at least in part upon the plurality of theoretical fair market prices computed by that computational module, wherein the identified band has a width that is less than a full range of the volatility values for which that computational module computed the plurality of theoretical fair market prices, wherein each next  
25 successive computational module in the pipeline is configured to operate on a plurality of volatility values within the band identified by the previous computational module in the pipeline.

36. The device as defined in any of the preceding claims 27-35  
30 wherein each computational module is configured to compute the theoretical fair market prices based on a Cox-Ross-Rubenstein (CRR) option pricing model.

37. The device as defined in any of the preceding claims 27-36  
35 wherein each computational module is further configured to compute the theoretical fair market prices using at least one look-up table

of precomputed terms for the CRR option pricing model, the look-up table being indexed by at least one of the option characteristic values.

5 38. The device as defined in any of the preceding claims 27-37 further comprising at least one reconfigurable logic device on which the pipeline of computational modules is deployed.

39. A computer-implemented method for determining an implied  
10 volatility for an option on an underlying financial instrument, the method comprising:

computing a plurality of theoretical fair market prices for the option, each theoretical fair market price being computed based at least in part upon a volatility value for the underlying  
15 financial instrument;

identifying a band of volatility values within which the theoretical fair market price for the option differs from an actual purchase price for the option by less than a predetermined threshold;

20 iteratively narrowing the band of volatility values and repeating the computing step for a plurality of volatility values within the narrowed band until one of the computed theoretical fair market prices differs from the actual purchase price by less than the predetermined threshold; and

25 defining the implied volatility for the option as the volatility value for which the computed theoretical fair market price differs from the actual purchase price by less than the predetermined threshold.

30 40. The method as defined in claim 39 further comprising:

for a first iteration, defining a plurality of volatility values such that a range of volatility values for the underlying financial instrument is subdivided into a plurality  $m$  of bands; and

35 wherein the computing step for the first iteration comprises computing the theoretical fair market prices for each of the defined volatility values.

41. The method as defined in any of the preceding claims 39-40 wherein the band identifying step comprises:

determining which one of the m bands has a computed  
5 theoretical fair market price at a first boundary and a computed theoretical fair market price at a second boundary such that the actual purchase price falls between the computed theoretical fair market price at the first boundary and the computed theoretical fair market price at the second boundary; and

10 selecting the determined band as the identified band.

42. The method as defined in any of the preceding claims 39-41 wherein the iteratively narrowing and repeating step comprises

for each iteration, (1) defining a plurality of volatility  
15 values within the selected band such that the selected band is subdivided into a plurality m of bands, (2) computing a plurality of theoretical fair market prices for the defined volatility values within the selected band, and (3) performing the band identifying step for the m bands of the selected band.

20 43. The method as defined in any of the preceding claims 39-42 wherein the value of m for all iterations is 4.

44. The method as defined in any of the preceding claims 39-43  
25 wherein the computing step comprises computing the plurality of theoretical fair market prices in parallel.

45. The method as defined in any of the preceding claims 39-44 wherein the computing step comprises computing the theoretical fair  
30 market prices based on a Cox-Ross-Rubinstein (CRR) option pricing model.

46. The method as defined in claim 45 wherein the computing step further comprises computing the plurality of theoretical fair market  
35 prices in parallel based on the CRR option pricing model.

47. The method as defined in claim 45 wherein the computing step further comprises computing each theoretical fair market price using a binomial tree having a plurality of CRR computational nodes, the binomial tree being of depth  $n$ , wherein the computational nodes of the same depth are configured to operate in parallel.

48. The method as defined in any of the preceding claims 39-47 further comprising:

performing the computing step with reconfigurable logic.

49. The method as defined in any of the preceding claims 39-48 further comprising:

performing at least the computing step with reconfigurable logic.

50. The method as defined in any of the preceding claims 39-49 further comprising:

performing all of the method steps with reconfigurable logic.

51. The method as defined in any of the preceding claims 39-50 further comprising:

feeding the implied volatility to a black box trading algorithm; and

automatically determining through the black box trading algorithm whether to execute a trade on the option based at least in part upon the implied volatility fed thereto.

52. The method as defined in any of the preceding claims 39-51 further comprising:

automatically generating an option quote with an auto-quote server based at least in part upon the defined implied volatility.

53. The method as defined in any of the preceding claims 39-52 further comprising:

receiving data describing the option within a financial market data stream, the option data comprising the actual purchase price

for the option, a current price of the financial instrument underlying the option, a strike price for the option, and a time to maturity for the option.

5 54. A computer-implemented method for determining an implied volatility for an option on an underlying financial instrument, the method comprising:

10 computing a plurality of theoretical fair market prices for the option, each of said computed theoretical fair market prices corresponding to an assumed volatility value for the underlying financial instrument;

15 comparing the computed theoretical fair market prices with an actual purchase price for the option to determine whether at least one of the computed theoretical fair market prices satisfies a predetermined condition and,

if the pre-determined condition is satisfied, then declaring the volatility value for which the computed theoretical fair market price approximates the offer price as the option's implied volatility, and,

20 if the predetermined condition is not satisfied, identifying at least one computed theoretical fair market price in relative close proximity to the actual purchase price and selecting a band of volatility values within which the actual purchase price is expected to be found based on the identified at least one computed  
25 theoretical fair market price, and within the selected band, computing a second plurality of theoretical fair market prices and repeating the comparing step.

30 55. The method of claim 54 wherein the step of selecting a band of volatility values includes selecting the band between two computed theoretical fair market prices surrounding the actual purchase price.

56. A device comprising:

35 an option pricing engine that is configured to process a financial market data stream to compute an implied volatility for an

option on an underlying financial instrument, wherein the option pricing engine is configured to iteratively subdivide a range of volatility values for the underlying financial instrument into a plurality  $m$  of bands and, for each band compute at least two  
5 theoretical fair market prices for the option based on at least two volatility values for the band;

wherein the option pricing engine is further configured to continue the iterative subdivision and the corresponding theoretical fair market computations until an implied volatility for the option  
10 is found as determined by at least one condition.

57. The device as defined in claim 56 wherein the option pricing engine comprises a pipeline of computational modules, each computational module being configured to compute a plurality of the  
15 theoretical fair market prices for a different iterative subdivision.

58. The device as defined in claim 57 wherein the option pricing engine is further configured to feed data describing the option into  
20 the first computational module in the pipeline and initialize the first computational module with a plurality of parameters that control how the theoretical fair market prices will be computed.

59. The device as defined in claim 58 wherein the option data  
25 comprises the actual purchase price for the option, a strike price for the option, a price for the underlying financial instrument, and a time to maturity for the option, and wherein the parameters comprise a plurality of volatility values for computing the theoretical fair market prices of the first iteration.

30

60. The device as defined in any of the preceding claims 58-59 wherein the parameters further comprise an interest rate related to the option.

35 61. The device as defined in any of the preceding claims 56-60 wherein the option pricing engine is further configured to compute

the plurality of volatility values for use in computing the theoretical fair market prices within each band for each iteration.

62. The device as defined in any of the preceding claims 57-61 wherein each computational module is configured to compute the plurality of theoretical fair market prices for the different volatility values in parallel.

63. The device as defined in any of the preceding claims 56-62 wherein the option pricing engine is further configured to identify which one of the m bands that the implied volatility resides in by determining the band for which the actual purchase price falls between the computed theoretical fair market prices for that band.

64. The device as defined in any of the preceding claims 57-63 wherein each computational module is configured to compute the plurality of theoretical fair market prices based on a Cox-Ross-Rubinstein (CRR) option pricing model.

65. The device as defined in claim 64 wherein each computational module further comprises a plurality of parallel CRR option pricing model computational units, each parallel CRR option pricing model computational unit being configured to receive a volatility value as an input parameter and compute a theoretical fair market price for the option based at least in part upon the received volatility value.

66. The device as defined in claim 65 wherein each parallel CRR option pricing model computational unit comprises a binomial tree of CRR computational nodes, the binomial tree comprising a plurality n of pipelined stages, wherein the computational nodes sharing a common pipelined stage in the tree are configured to operate in parallel, and wherein the pipelined stages are arranged such that the computational nodes at a lower level of the tree feed their outputs to the computational nodes at a next higher level of the tree, and wherein the computational node at a root of the tree is

configured to output the computed theoretical fair market price for the input volatility value.

67. The device as defined in any of the preceding claims 65-66 wherein the plurality of parallel CRR option pricing units are deployed on at least one reconfigurable logic device.

68. The device as defined in any of the preceding claims 57-60, 62, or 64-67 wherein the computational modules are deployed on at least one reconfigurable logic device.

69. The device as defined in claim 59 wherein the plurality of volatility values used to initialize the first computational module in the pipeline comprise a plurality of volatility values within a predetermined range of volatility values.

70. The device as defined in claim 59 wherein the option pricing engine is further configured to select the plurality of volatility values used to initialize the first computational module in the pipeline based on a plurality of historical volatility values for the financial instrument underlying the option.

71. The device as defined in any of the preceding claims 56-70 wherein the at least one condition comprises a theoretical fair market price convergence property.

72. The device as defined in any of the preceding claims 56-71 wherein the at least one condition comprises a volatility convergence property.

73. The device as defined in any of the preceding claims 56-72 wherein the at least one condition comprises an iteration maximum.

74. The device as defined in any of the preceding claims 56-73 wherein the at least one condition comprises an adjustable condition.

75. The device as defined in any of the preceding claims 56-74 wherein the option pricing engine is deployed on at least one reconfigurable logic device.

5

76. The device as defined in any of the preceding claims 56-75 wherein the option pricing engine is deployed across a plurality of parallel co-processors.

10 77. A computer-implemented method for computing an implied volatility for an option on an underlying financial instrument, the method comprising:

receiving data describing an option on an underlying financial instrument; and

15 performing an iterative m-ary search within a range of volatility values for the underlying financial instrument using the option data to thereby compute the implied volatility for the option.

20 78. The method of claim 77 further comprising:

performing the iterative m-ary search with reconfigurable logic.

79. A computer-implemented method for determining an implied  
25 volatility for an option on an underlying financial instrument, the method comprising:

computing a plurality of theoretical fair market prices for a plurality of different volatility values;

30 choosing at least one computed theoretical fair market price relatively close to an actual purchase price for the option;

iteratively computing a plurality of theoretical fair market prices in correspondingly narrowing bands of volatility values until a predetermined condition is satisfied; and

35 in response to satisfaction of the predetermined condition, defining the implied volatility for the option based at least in

part on a volatility value used to compute a theoretical fair market price when the predetermined condition was satisfied.

80. The method of claim 79 further comprising:

5 performing the method steps with reconfigurable logic.

81. A computer-implemented method for computing a value of interest in relation to a financial instrument, the method comprising:

10 receiving financial market data that is descriptive of a financial instrument, wherein the financial instrument is associated with at least a first parameter and a second parameter, wherein the first parameter is a monotonically increasing or decreasing function of the second parameter;

15 iteratively computing a value of interest for the financial instrument by:

(1) defining a plurality of second parameter values, wherein the plurality of second parameter values together define a range of second parameter values that has been subdivided into a plurality *m* of bands;

20 (2) for each defined second parameter value, computing a value for the first parameter; and

(3) for each computed first parameter value, computing a value that is indicative of how closely the first parameter value and the received financial market data value match each other;

25 (4) for each computed indicative value, comparing the computed indicative value with a predetermined threshold to thereby determine whether any additional iterations are needed;

30 (5) responsive to a determination that no additional iterations are needed, outputting as the value of interest the first parameter value used to compute the second parameter value for which the comparing step resulted in a determination that no further iterations are needed, and responsive to a determination that an additional iteration is needed,

35

repeating steps (1)-(5) for a next iteration, wherein for the next iteration the range of second parameter values in the defining step is the band for which the received financial market data value falls between the two first parameter values computed for that band.

82. The method as defined in claim 81 wherein the range for a first iteration is a full range of second parameter values for the function.

83. The method as defined in any of the preceding claims 81-82 wherein the financial instrument comprises a derivative financial instrument.

84. The method as defined in claim 83 wherein the derivative financial instrument comprises an option on an underlying financial instrument.

85. The method as defined in any of the preceding claims 81-84 wherein the received financial market data value comprises a price for the option, wherein the first parameter comprises a theoretical fair purchase price for the option, wherein the second parameter comprises a volatility for the underlying financial instrument, and wherein the value of interest comprises an implied volatility for the option.

86. The method as defined in any of the preceding claims 81-85 further comprising performing the iteratively computing step with reconfigurable logic.

87. The method as defined in any of the preceding claims 81-86 wherein the first parameter is a monotonically strictly increasing or strictly decreasing function of the second parameter.

88. A device for computing a price for an option, the device comprising:

an options pricing engine configured to process data describing an option through an option pricing model to thereby compute a theoretical fair market price for the option;

wherein the options pricing engine comprises a plurality of  
5 pipelined stages, the plurality of pipelined stages comprising a plurality of parallel computational nodes, each of said nodes being configured to compute a step of a stepwise option pricing model, and wherein said computational nodes are cascaded such that each successive pipelined stage is configured to receive as its input the  
10 output from at least two computational nodes in the immediately preceding pipelined stage, wherein the output of the last pipelined stage comprises the computed theoretical fair market price for the option.

15 89. The device as defined in claim 88 wherein the option pricing model comprises a Cox, Ross and Rubinstein (CRR) option pricing model.

90. The device as defined in any of the preceding claims 88-89  
20 wherein the pipelined stages are arranged in a binomial tree of computational nodes, wherein the tree is of depth n.

91. The device as defined in claim 90 wherein the options pricing engine is configured to partition the binomial tree across a  
25 plurality of processing resources.

92. The device as defined in claim 91 wherein the plurality of processing resources comprise processing resources of different types.  
30

93. The device as defined in any of the preceding claims 91-92 wherein at least one of the processing resources comprises a reconfigurable logic device.

35 94. The device as defined in any of the preceding claims 88-93 wherein the pipelined stages are configured to compute a plurality

of terminal financial instrument prices for the option pricing model based on an upward multiplicative factor and a downward multiplicative factor.

5 95. The device as defined in any of the preceding claims 88-95 wherein the options pricing engine further comprises a second set of the pipelined stages, wherein the second set of pipelined stages are configured to compute a plurality of probability terms for the option pricing model that are used in the theoretical fair market  
10 price computation.

96. The device as defined in claim 95 wherein the two sets of pipelined stages are arranged in parallel with each other.

15 97. The device as defined in any of the preceding claims 95-96 wherein the two sets of pipelined stages feed their respective outputs of terminal financial instrument prices and probability terms to a combinatorial logic stage that is configured to compute the theoretical fair market price based at least in part upon the  
20 terminal financial instrument prices and the probability terms.

98. The device as defined in any of the preceding claims 88-97 wherein the options pricing engine is further configured to perform at least a portion of the theoretical fair market price computation  
25 with reconfigurable logic.

99. A computer-implemented method comprising:

receiving data describing an option on an underlying financial instrument; and

30 computing a theoretical fair market price for the option using a plurality of pipelined stages, the plurality of pipelined stages comprising a plurality of parallel computational nodes, each of said nodes being configured to compute a step of a stepwise option pricing model, and wherein said computational nodes are cascaded  
35 such that each successive pipelined stage receives as its input the output from at least two computational nodes in the immediately

preceding pipelined stage, wherein the output of the last pipelined stage comprises the computed theoretical fair market price for the option.

5 100. The method as defined in claim 99 wherein the option pricing model comprises a Cox, Ross and Rubinstein (CRR) option pricing model.

10 101. The method as defined in any of the preceding claims 99-100 wherein the computing step further comprises providing a binomial tree of computational units, wherein the tree is of depth n.

15 102. The method as defined in claim 101 further comprising partitioning the binomial tree across a plurality of processing resources.

20 103. The method as defined in claim 102 wherein the plurality of processing resources comprise processing resources of different types.

104. The method as defined in any of the preceding claims 102-103 wherein at least one of the processing resources comprises a reconfigurable logic device.

25 105. The method as defined in any of the preceding claims 99-104 wherein the computing step further comprises computing a plurality of terminal financial instrument prices for the option pricing model with the pipelined stages based on an upward multiplicative factor and a downward multiplicative factor.

30 106. The method as defined in any of the preceding claims 99-105 wherein the computing step further comprises computing a plurality of probability terms for the option pricing model with the pipelined stages.

35

107. The method as defined in claim 106 wherein the computing step further comprises computing the terminal financial instrument prices and the probability terms in parallel.

5 108. The method as defined in any of the preceding claims 106-107 wherein the computing step further comprises feeding the terminal financial instrument prices and the probability terms to a combinatorial logic stage that is configured to compute the theoretical fair market price based at least in part upon the  
10 terminal financial instrument prices and the probability terms.

109. The method as defined in any of the preceding claims 99-108 further comprising performing the computing step in reconfigurable logic.

15

110. The method as defined in any of the preceding claims 99-109 further comprising:

feeding the computed theoretical fair market price to a black box trading algorithm; and

20

automatically determining through the black box trading algorithm whether to execute a trade on the option based at least in part upon the theoretical fair market price fed thereto.

111. The method as defined in any of the preceding claims 99-110 further comprising:

25

feeding the computed theoretical fair market price to an auto-quote server; and

automatically generating through the auto-quote server a price for an option quote based at least in part upon the theoretical fair  
30 market price fed thereto.

112. A device for computing a theoretical fair market price for a European option on an underlying financial instrument, the device comprising:

35

an options pricing engine that is configured to retrieve from a lookup table a plurality of precomputed terms for use in computing

the theoretical fair market price, wherein the options pricing engine is configured to retrieve the plurality of precomputed terms using an index based at least in part upon a volatility for the underlying financial instrument.

5

113. The device as defined in claim 112 wherein the options pricing engine comprises:

a first lookup table in which a plurality of volatility values for a plurality of different financial instruments are stored,

10 wherein the first lookup table is indexed by financial instrument;

a second lookup table in which the plurality of precomputed terms are stored, wherein the second lookup table is indexed by a volatility for the underlying financial instrument and a time to maturity for the option;

15 a lookup unit that is configured to process data describing the option, access the first lookup table based on an identity of the underlying financial instrument for the option to thereby retrieve the volatility value for that underlying financial instrument, access the second lookup table based on the retrieved volatility value and a time to maturity for the option to thereby  
20 retrieve the plurality of precomputed terms applicable to that option; and

25 a combinatorial logic stage that is configured to compute the theoretical fair market price according to an option pricing model based at least in part upon the data describing the option, the retrieved volatility value and the retrieved precomputed terms.

30 114. The device as defined in claim 113 wherein the option pricing model comprises a Cox, Ross and Rubinstein (CRR) option pricing model for computing a theoretical fair market price for a European option.

35 115. The device as defined in any of the preceding claims 113-114 wherein at least the combinatorial logic stage is implemented in reconfigurable logic.

116. A computer-implemented method for computing a theoretical fair market price for a European option on an underlying financial instrument, the method comprising:

retrieving, based at least in part upon a volatility for the underlying financial instrument, a plurality of precomputed terms from a lookup table for use in computing the theoretical fair market price; and

computing the theoretical fair market price for the option based at least in part upon the retrieved precomputed terms.

117. The method as defined in claim 116 further comprising:

receiving data describing the option, the data comprising an identity of the underlying financial instrument and a time to maturity for the financial instrument;

retrieving the volatility from another lookup table based on the identity of the underlying financial instrument, the another lookup table storing a plurality of volatility values for a plurality of different financial instruments, wherein the another lookup table is indexed by financial instrument; and

wherein the precomputed terms retrieving step further comprises retrieving the precomputed terms based at least in part upon the retrieved volatility and the time to maturity for the option.

118. The method of claim 117 wherein the option pricing model comprises a Cox, Ross and Rubinstein (CRR) option pricing model for European options.

119. The method as defined in any of the preceding claims 116-118 further comprising performing at least the computing step with reconfigurable logic.

120. The method as defined in any of the preceding claims 116-119 further comprising:

recomputing the precomputed terms according to at least one of the group consisting of a periodic basis and a triggered basis.

121. The method as defined in any of the preceding claims 116-120 further comprising:

5       feeding the computed theoretical fair market price to a black box trading algorithm; and

          automatically determining through the black box trading algorithm whether to execute a trade on the option based at least in part upon the theoretical fair market price fed thereto.

10 122. The method as defined in any of the preceding claims 116-121 further comprising:

          feeding the computed theoretical fair market price to an auto-quote server; and

15       automatically generating through the auto-quote server a price for an option quote based at least in part upon the theoretical fair market price fed thereto.

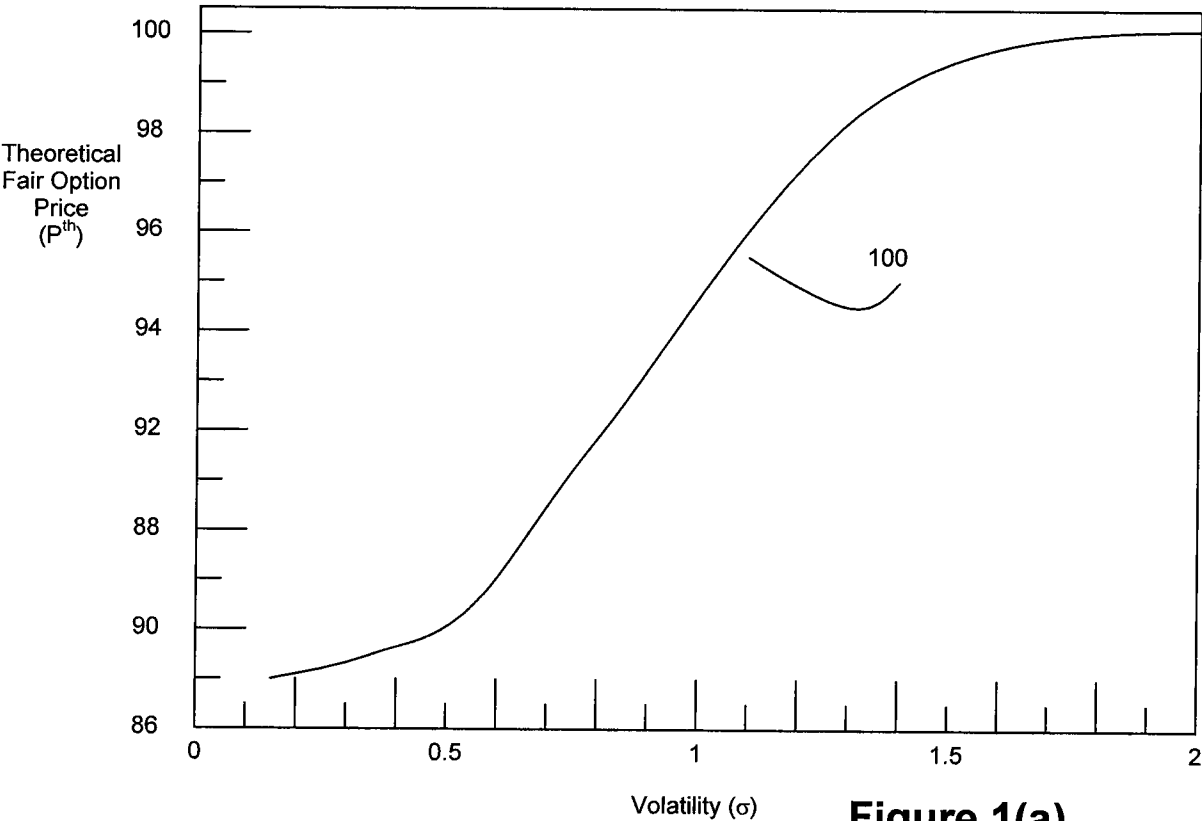


Figure 1(a)

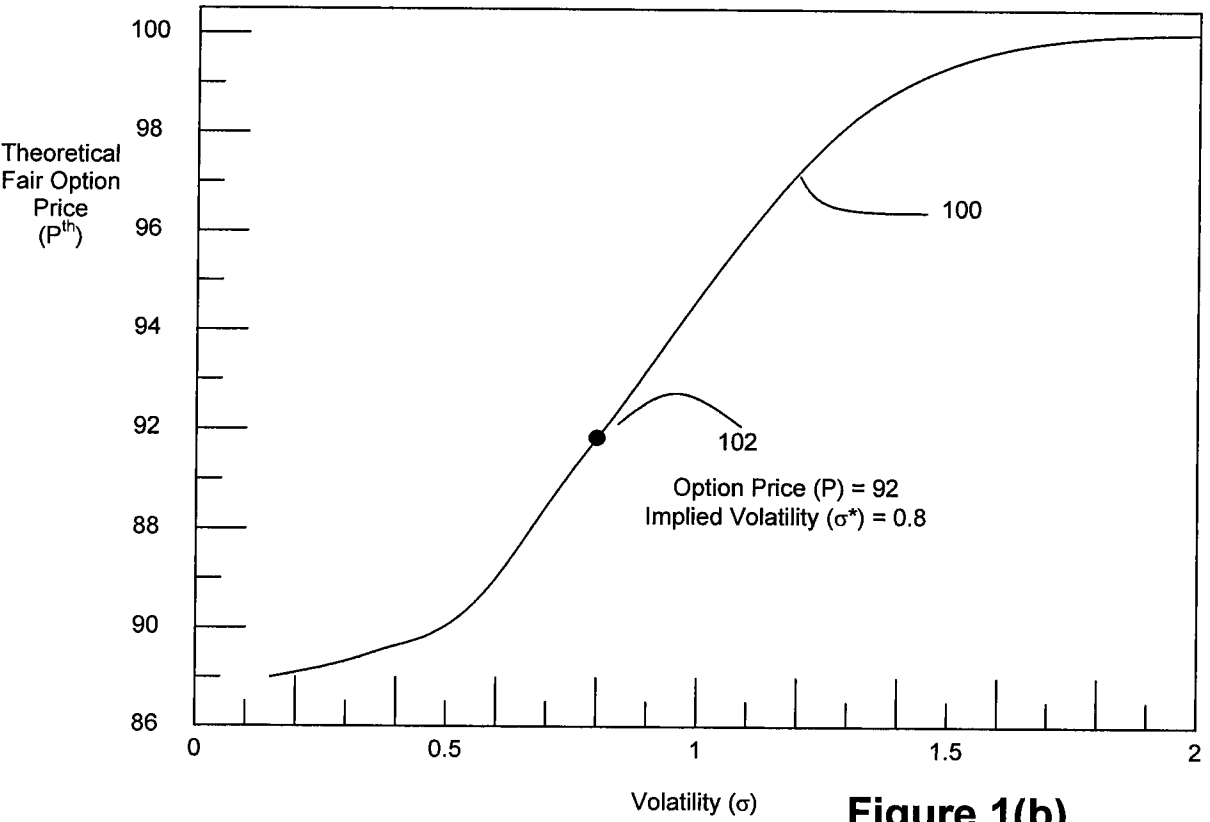
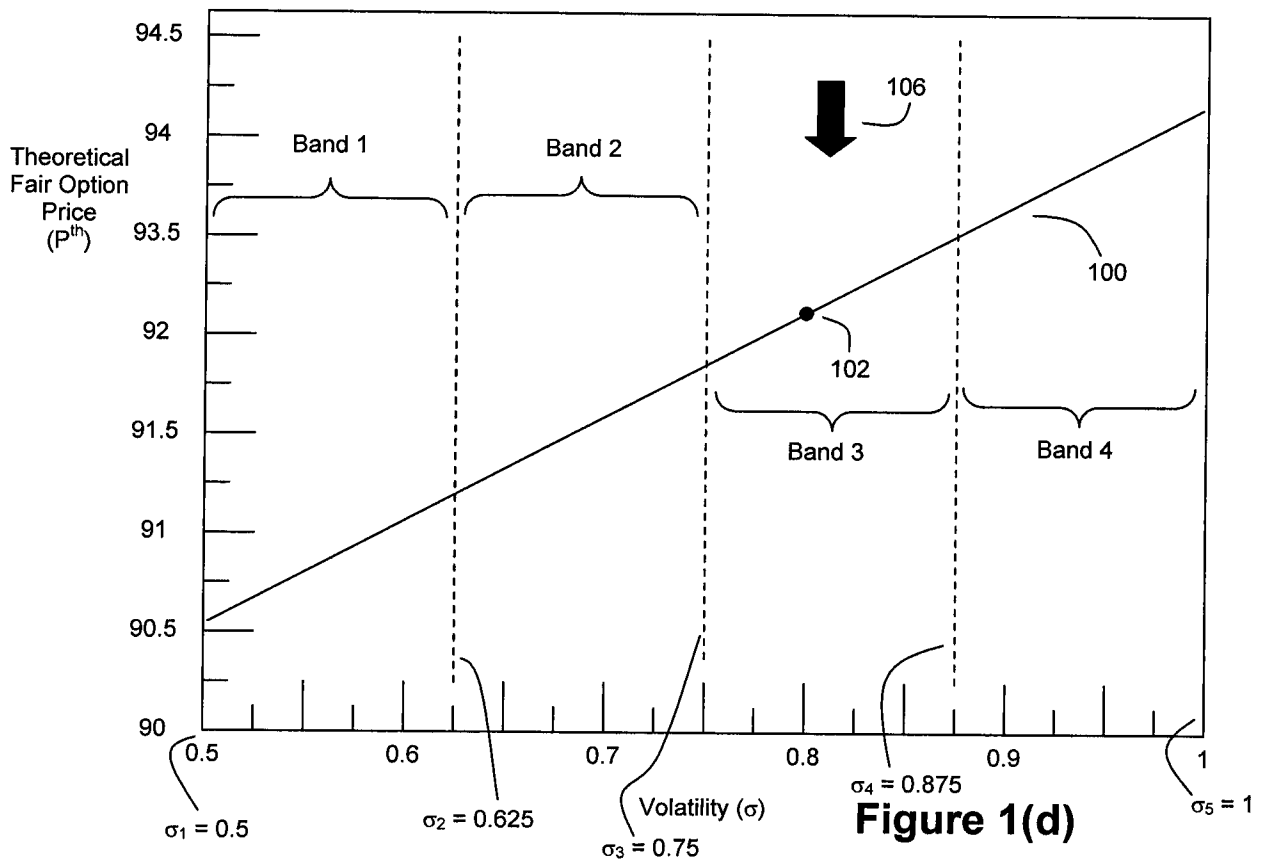
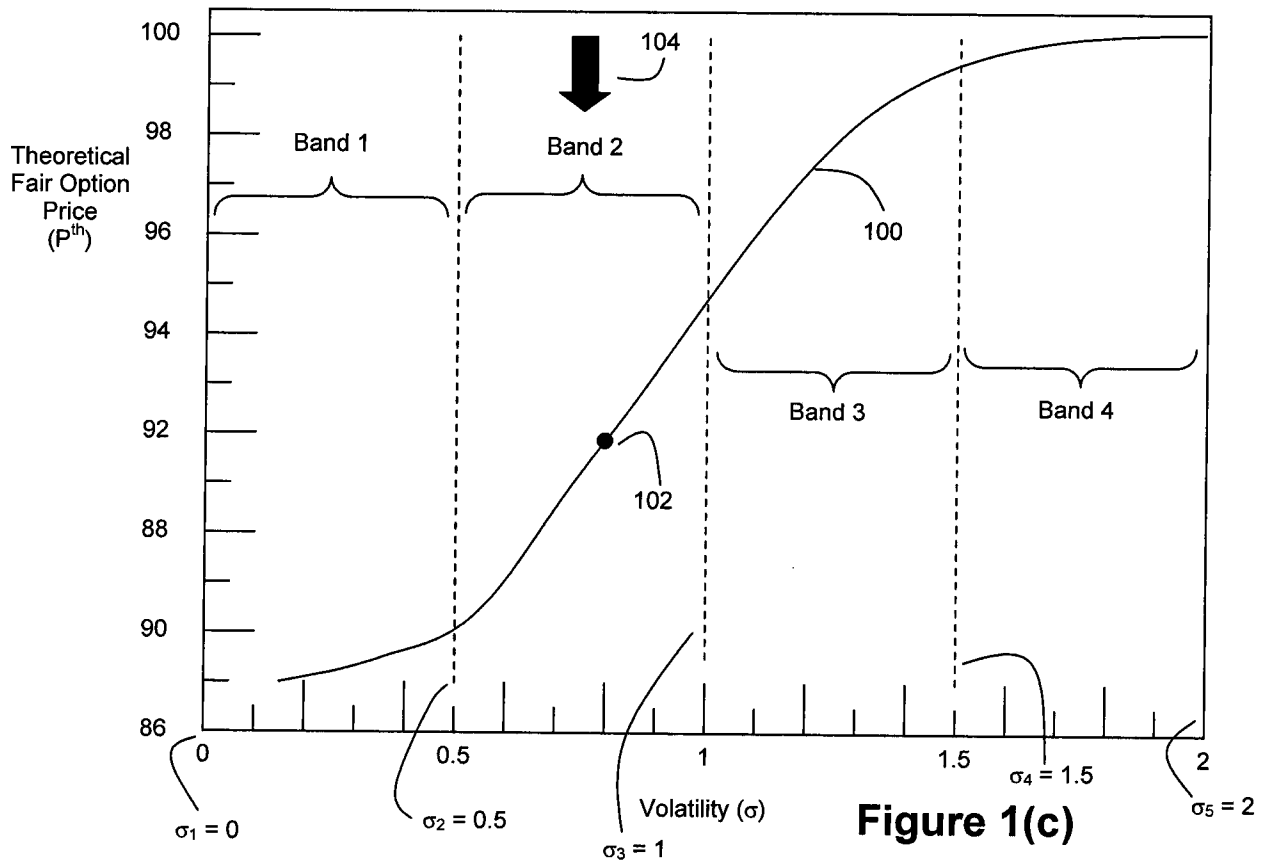


Figure 1(b)



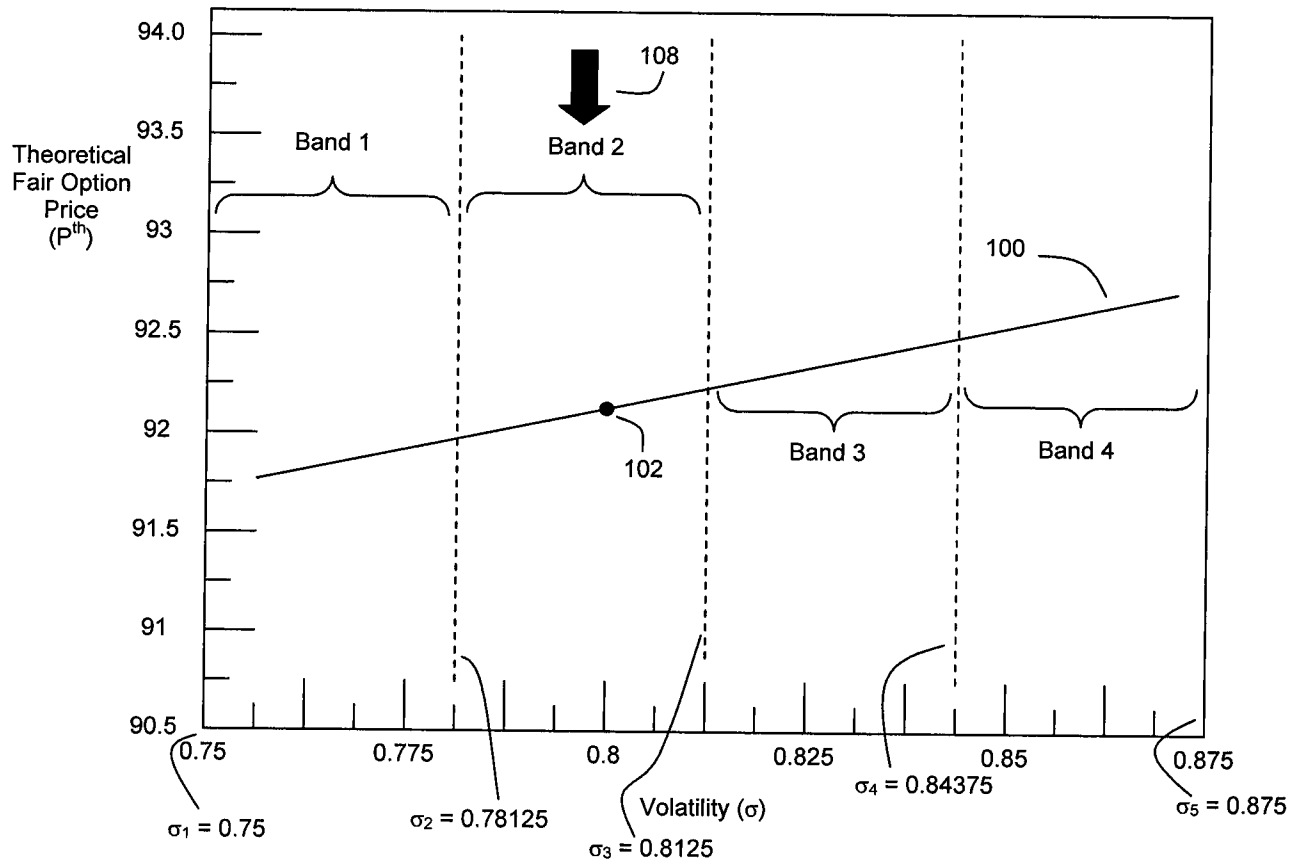


Figure 1(e)

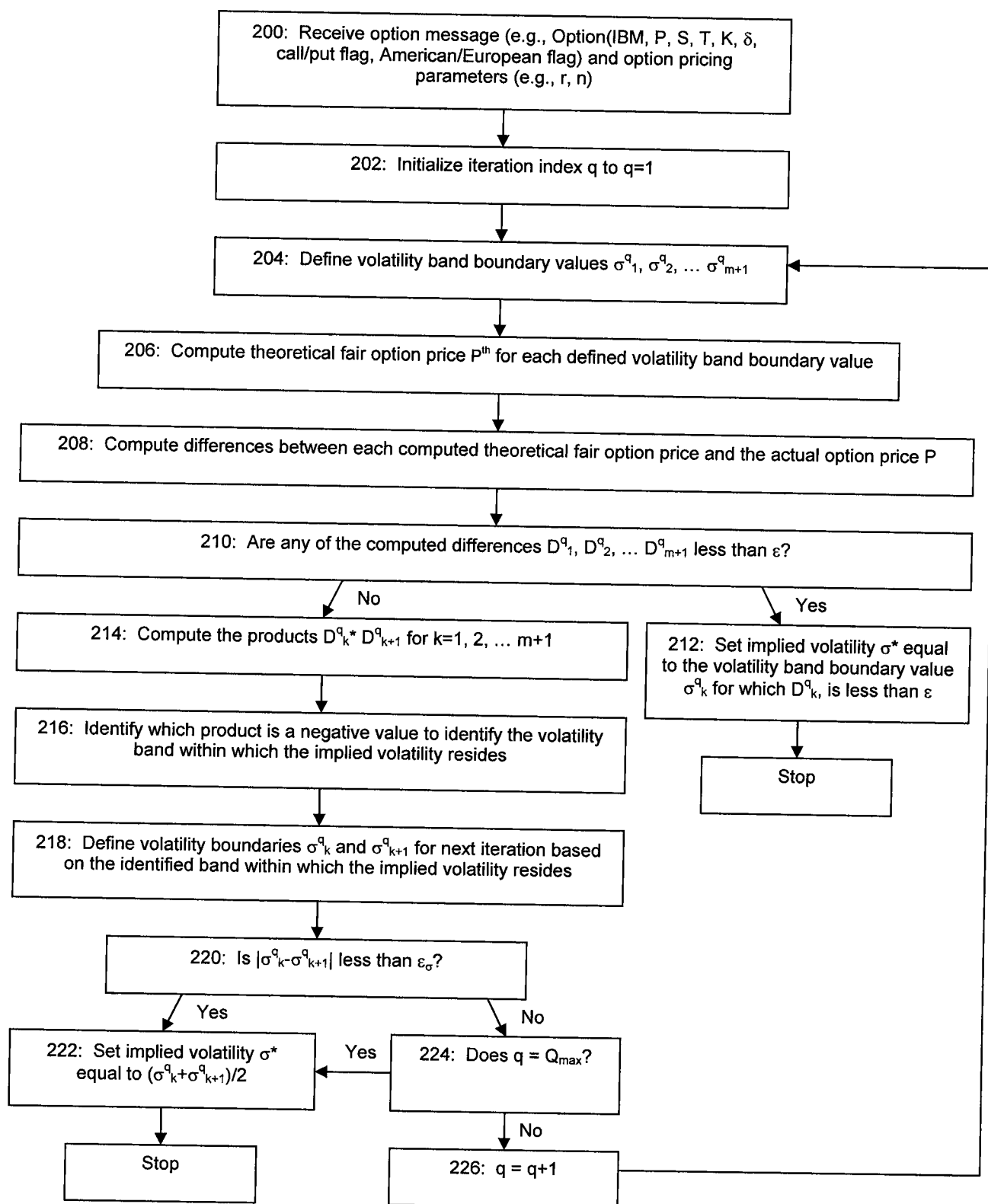


Figure 2(a)

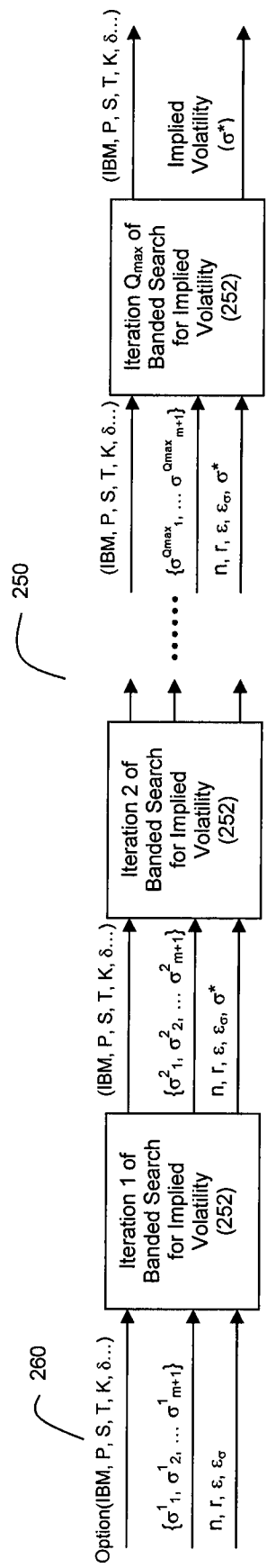


Figure 2(b)

260

|                      |                         |                                                     |                                  |                        |                              |                                     |
|----------------------|-------------------------|-----------------------------------------------------|----------------------------------|------------------------|------------------------------|-------------------------------------|
| Option Name<br>(262) | Option Price P<br>(264) | Underlying Financial<br>Instrument Price S<br>(266) | Time to Maturity<br>$T$<br>(268) | Call/Put Flag<br>(270) | Strike Price<br>$K$<br>(272) | Dividend Yield<br>$\delta$<br>(274) |
|----------------------|-------------------------|-----------------------------------------------------|----------------------------------|------------------------|------------------------------|-------------------------------------|

Figure 2(c)

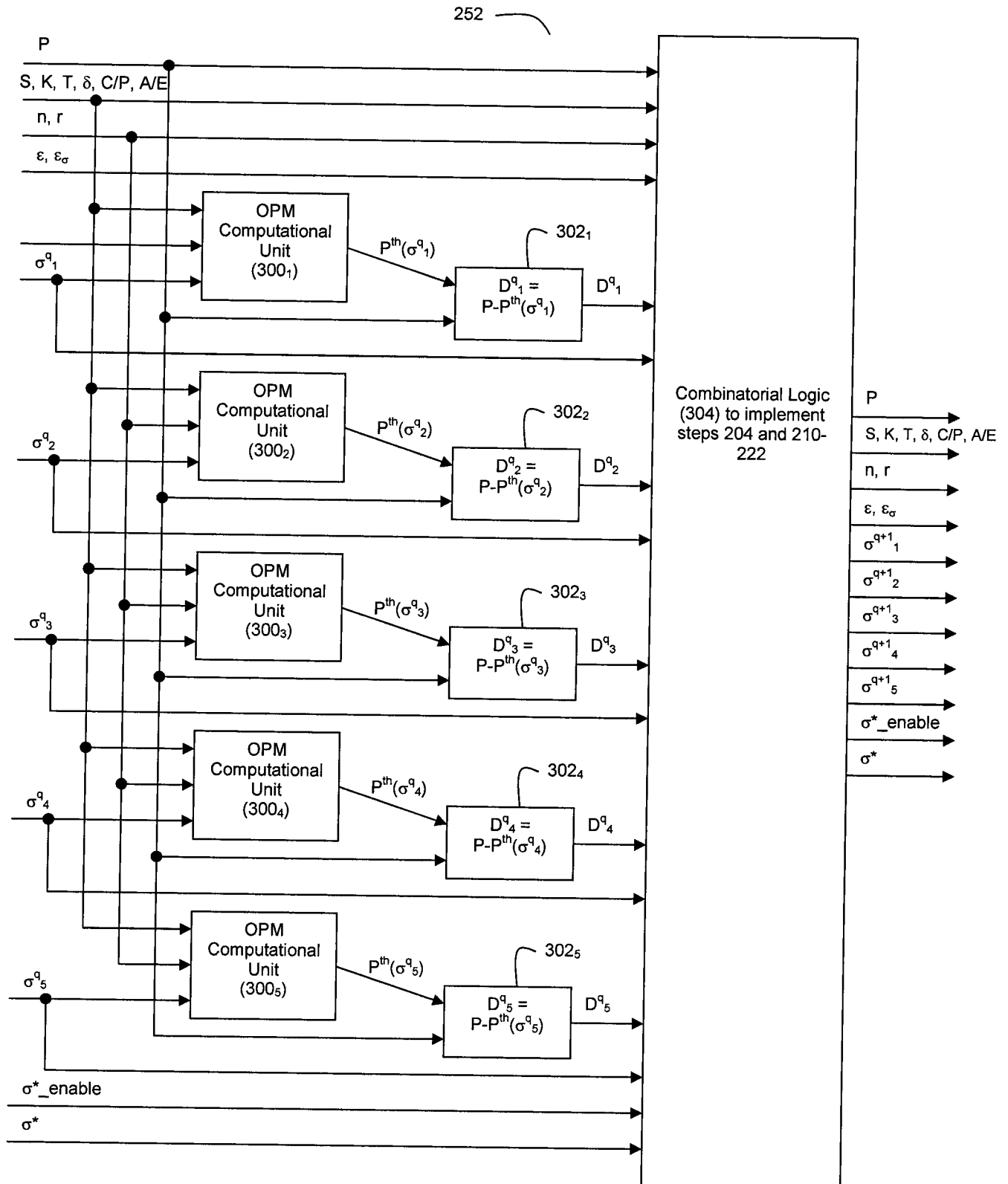


Figure 3

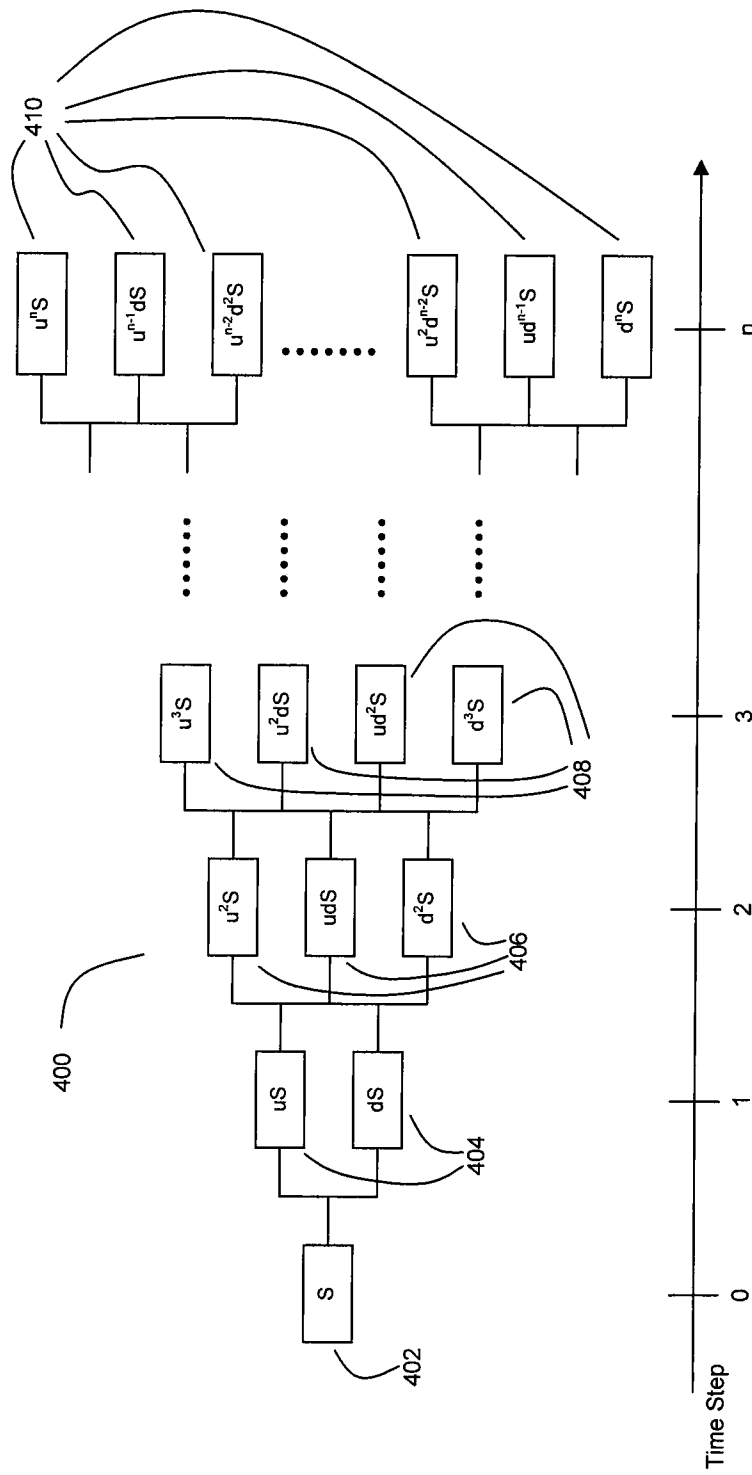


Figure 4

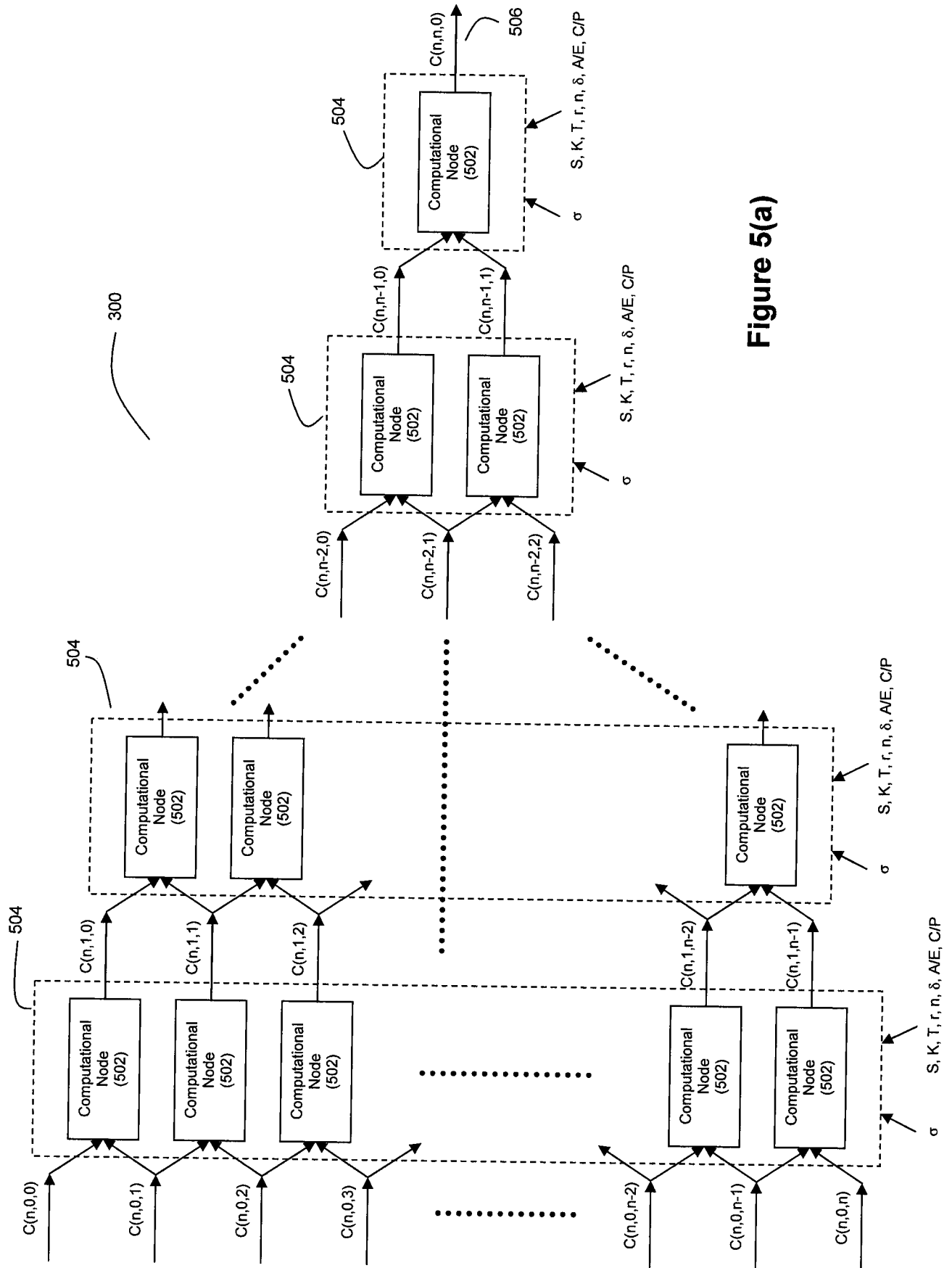


Figure 5(a)

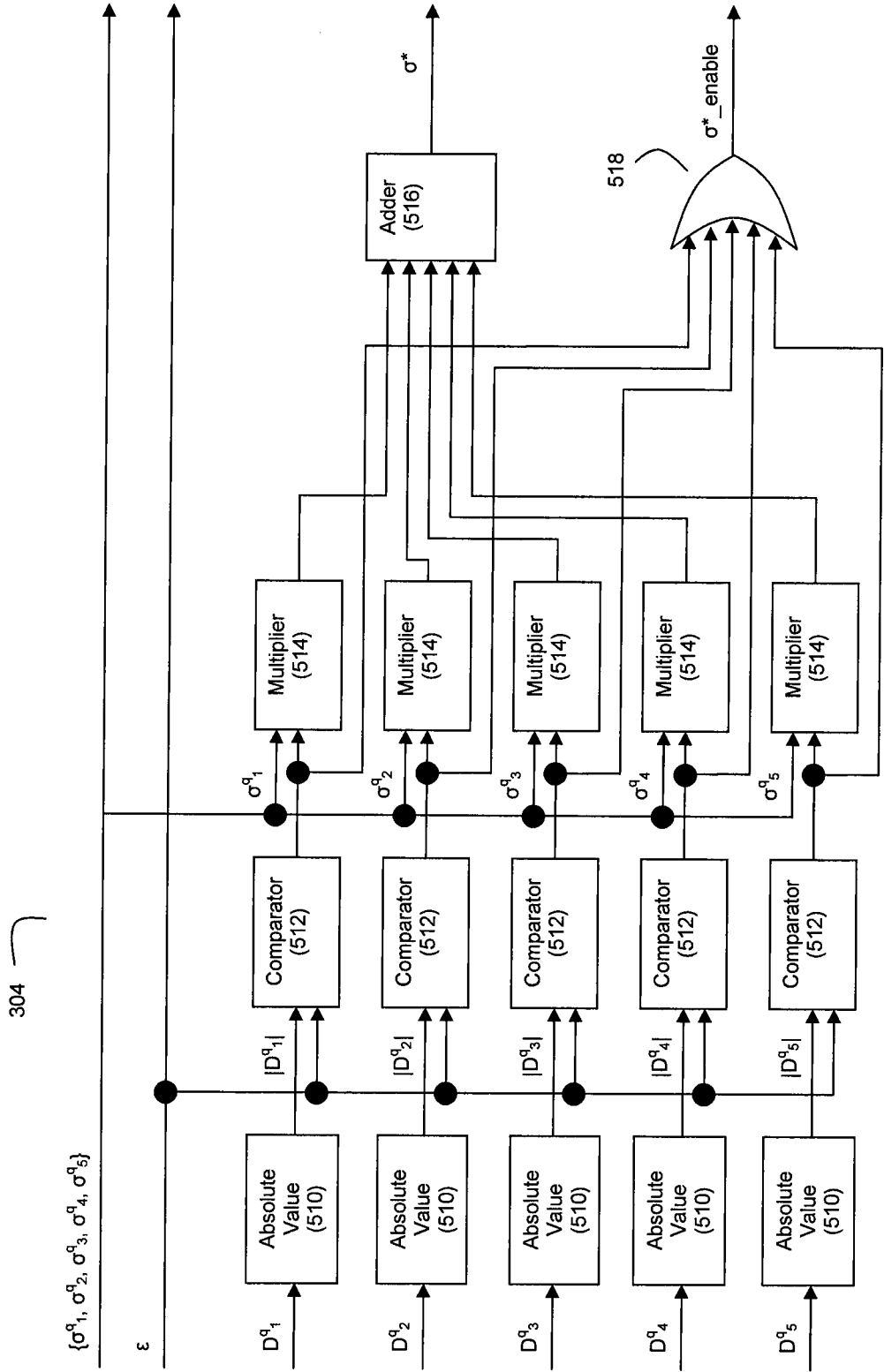


Figure 5(b)



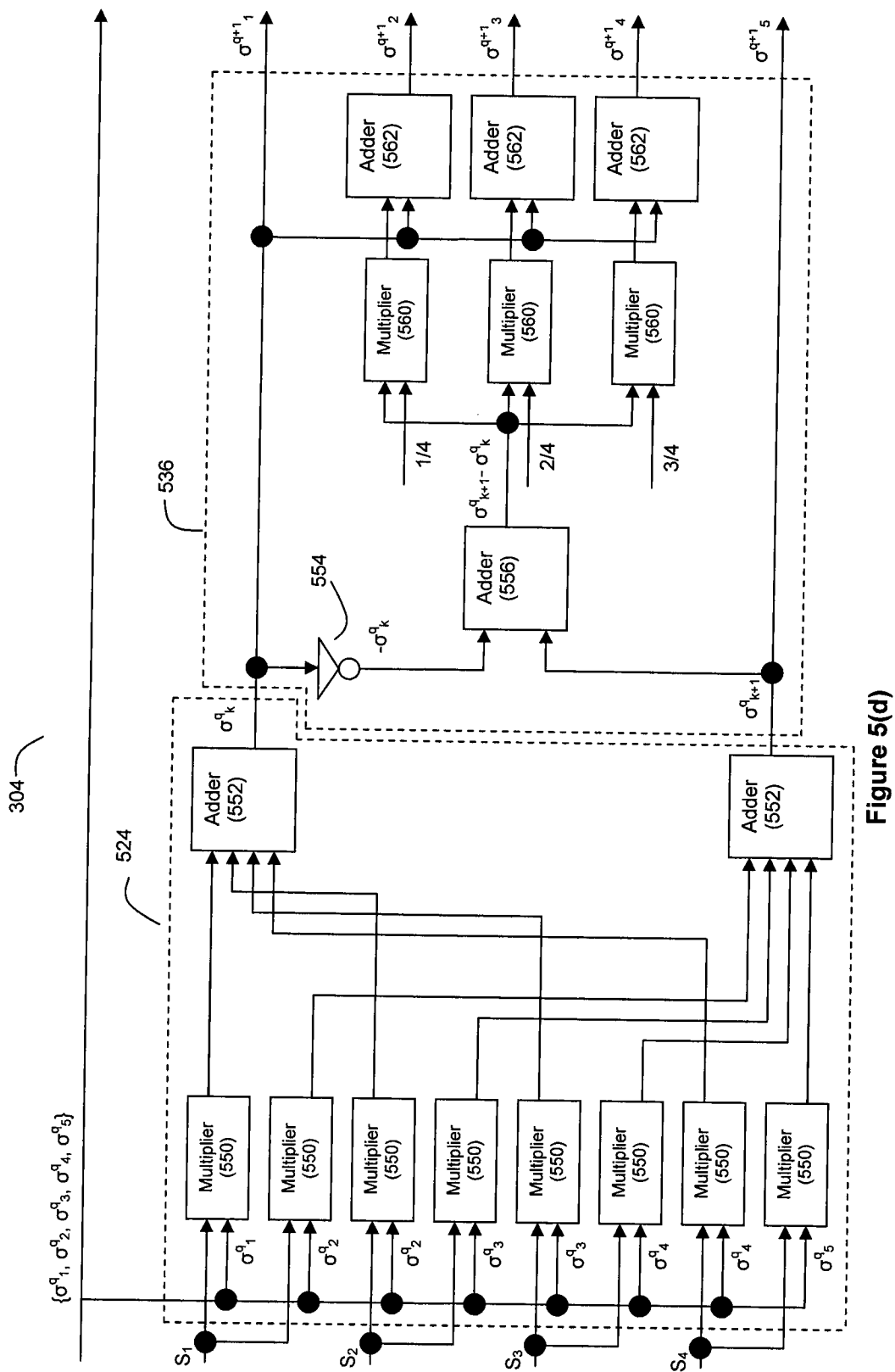


Figure 5(d)

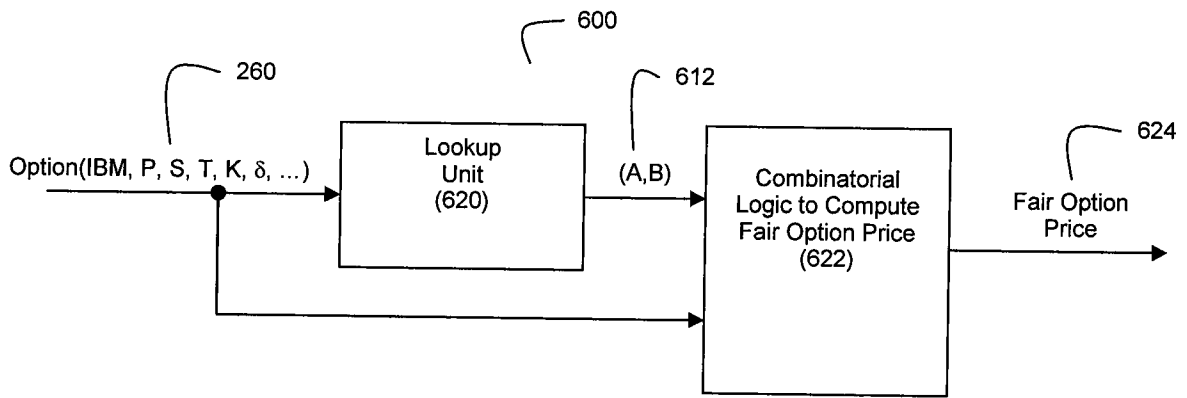


Figure 6(a)

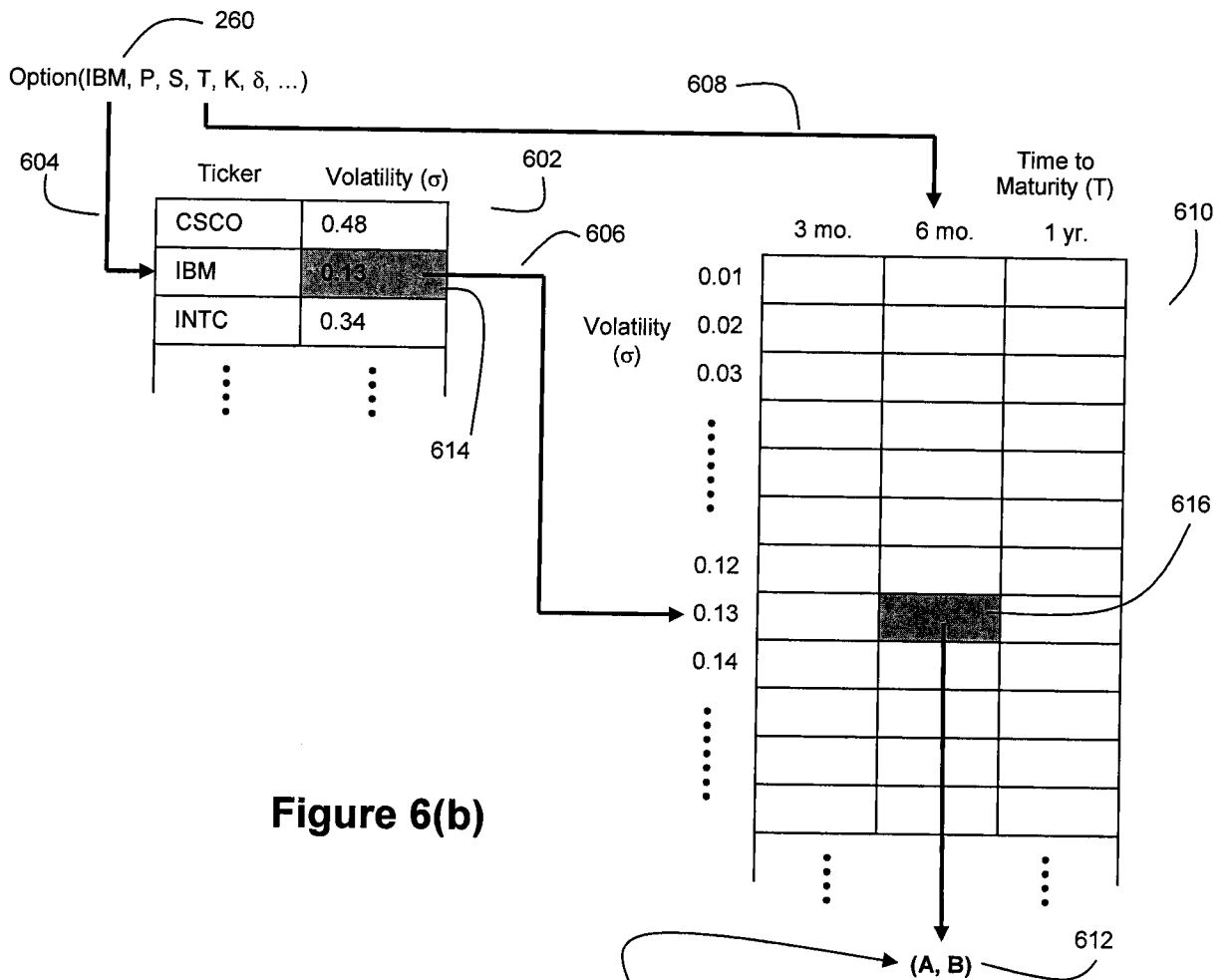


Figure 6(b)

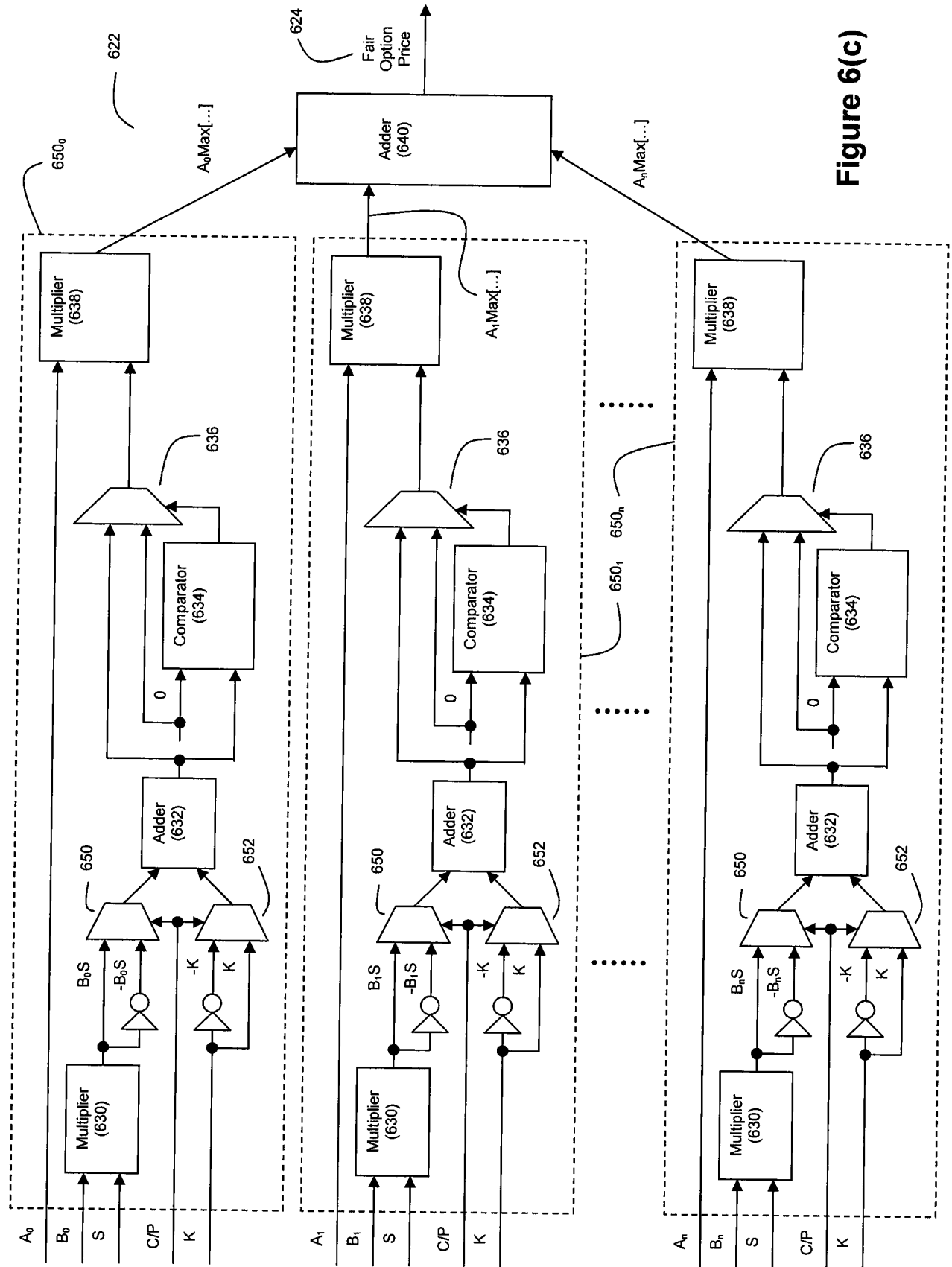
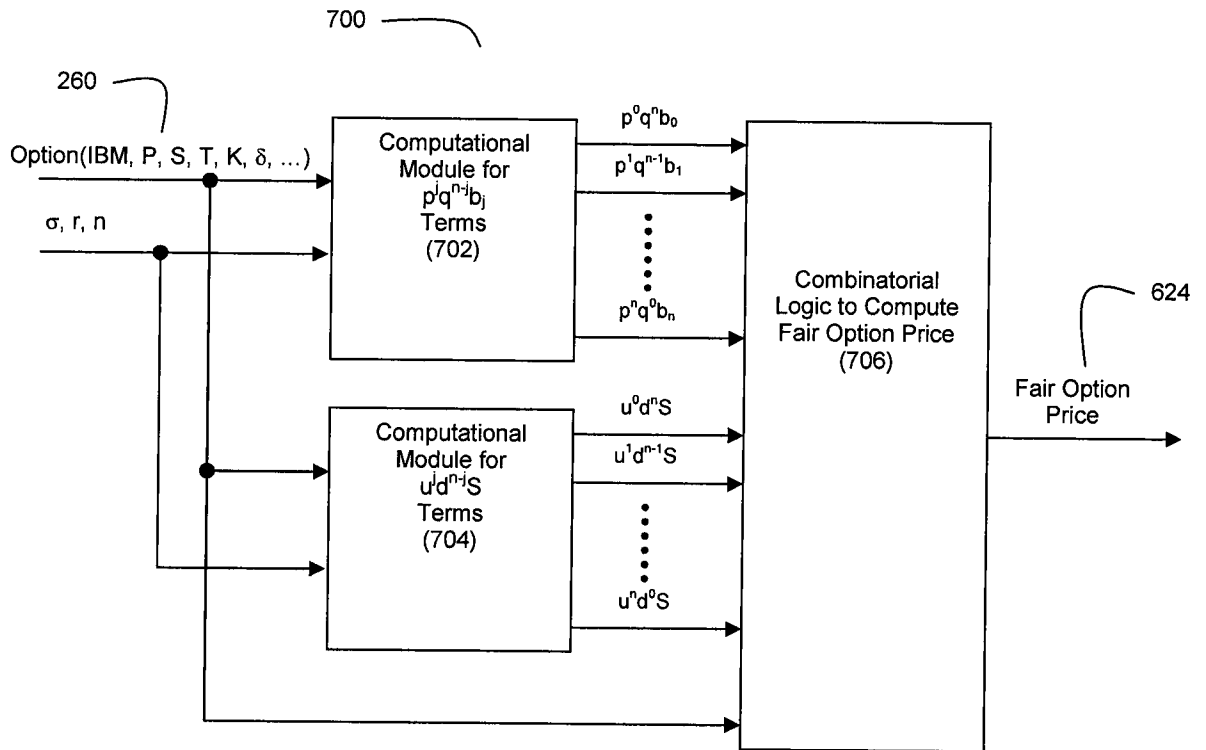


Figure 6(c)

**Figure 7**

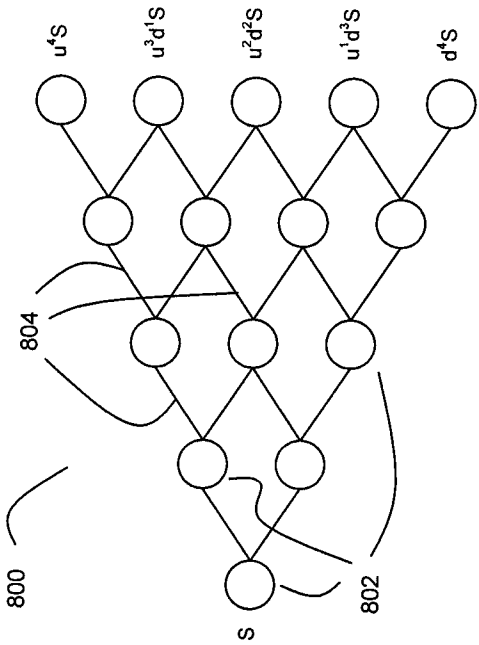


Figure 8(a)

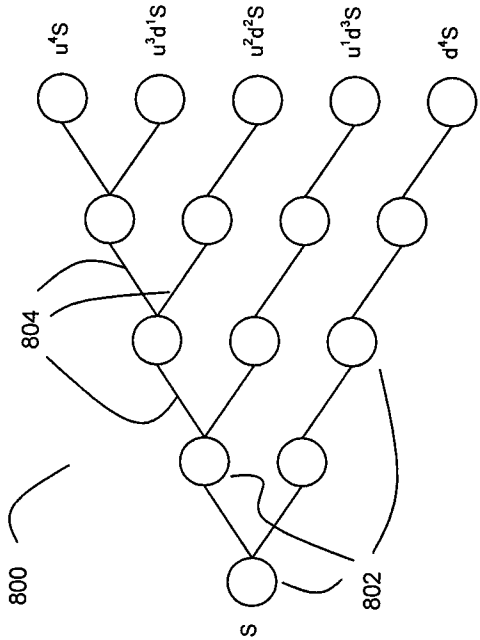


Figure 8(b)

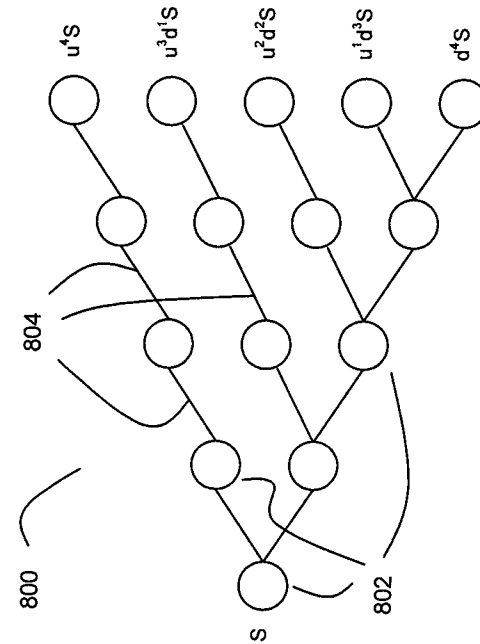


Figure 8(c)

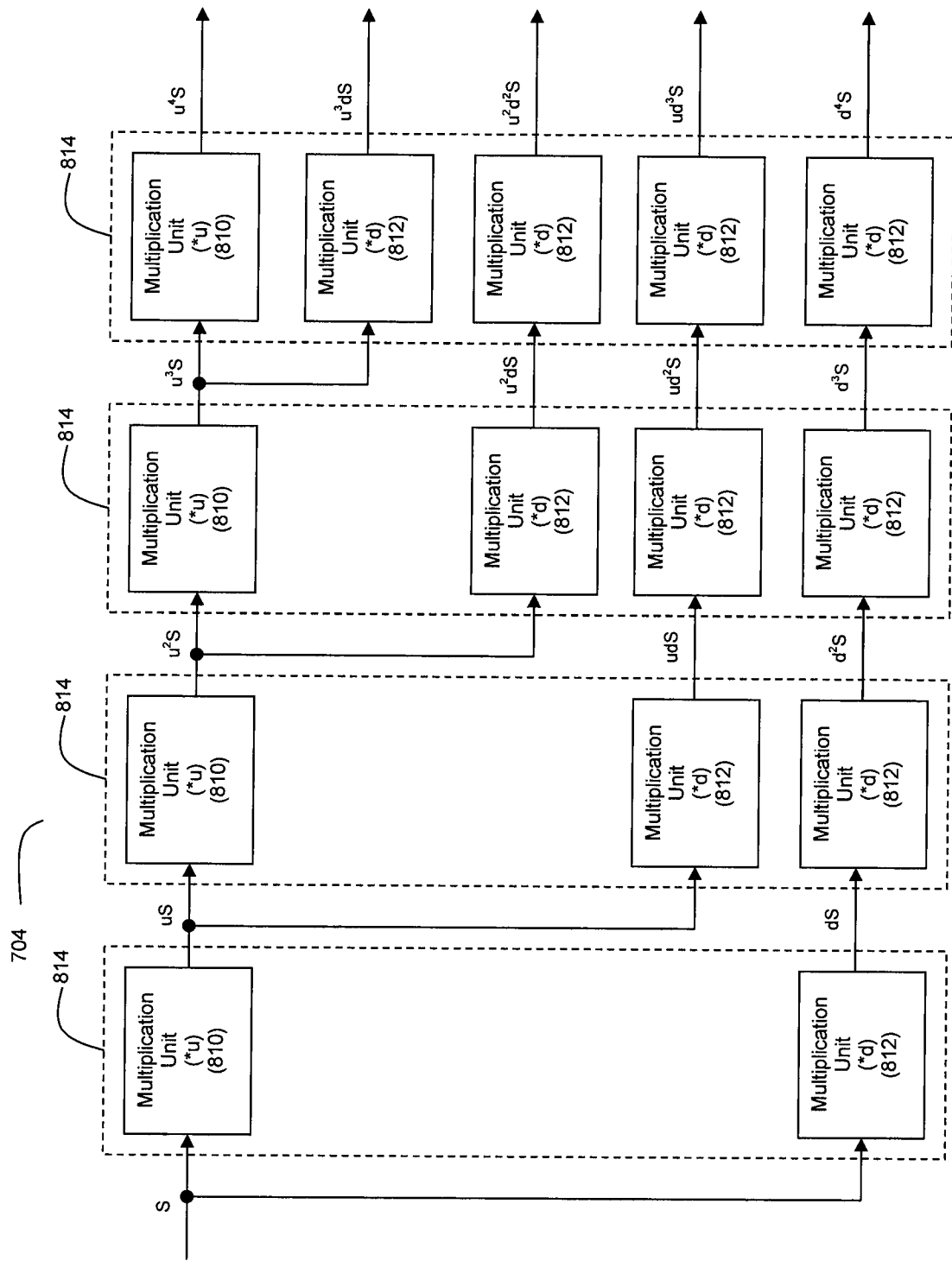


Figure 8(d)

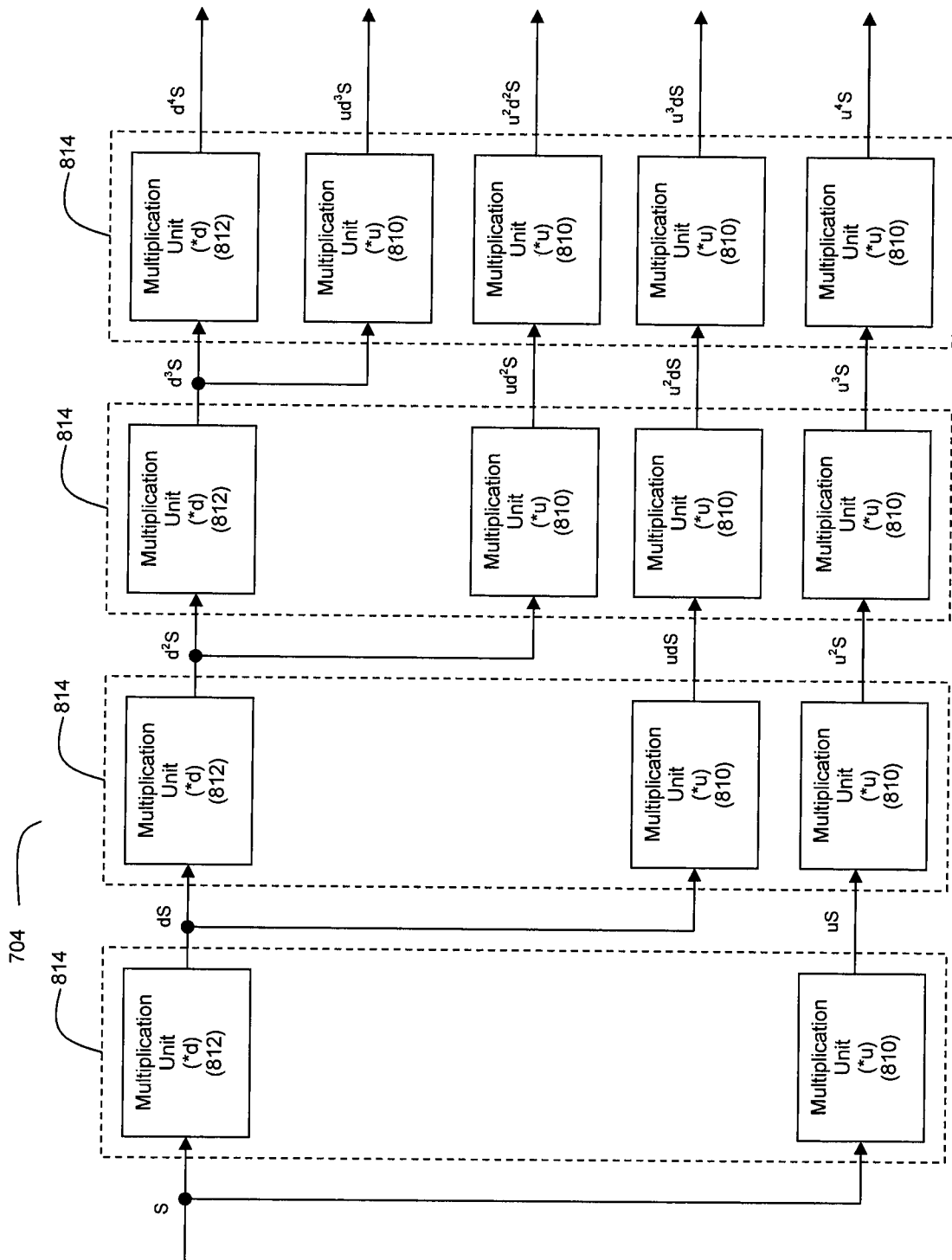


Figure 8(e)

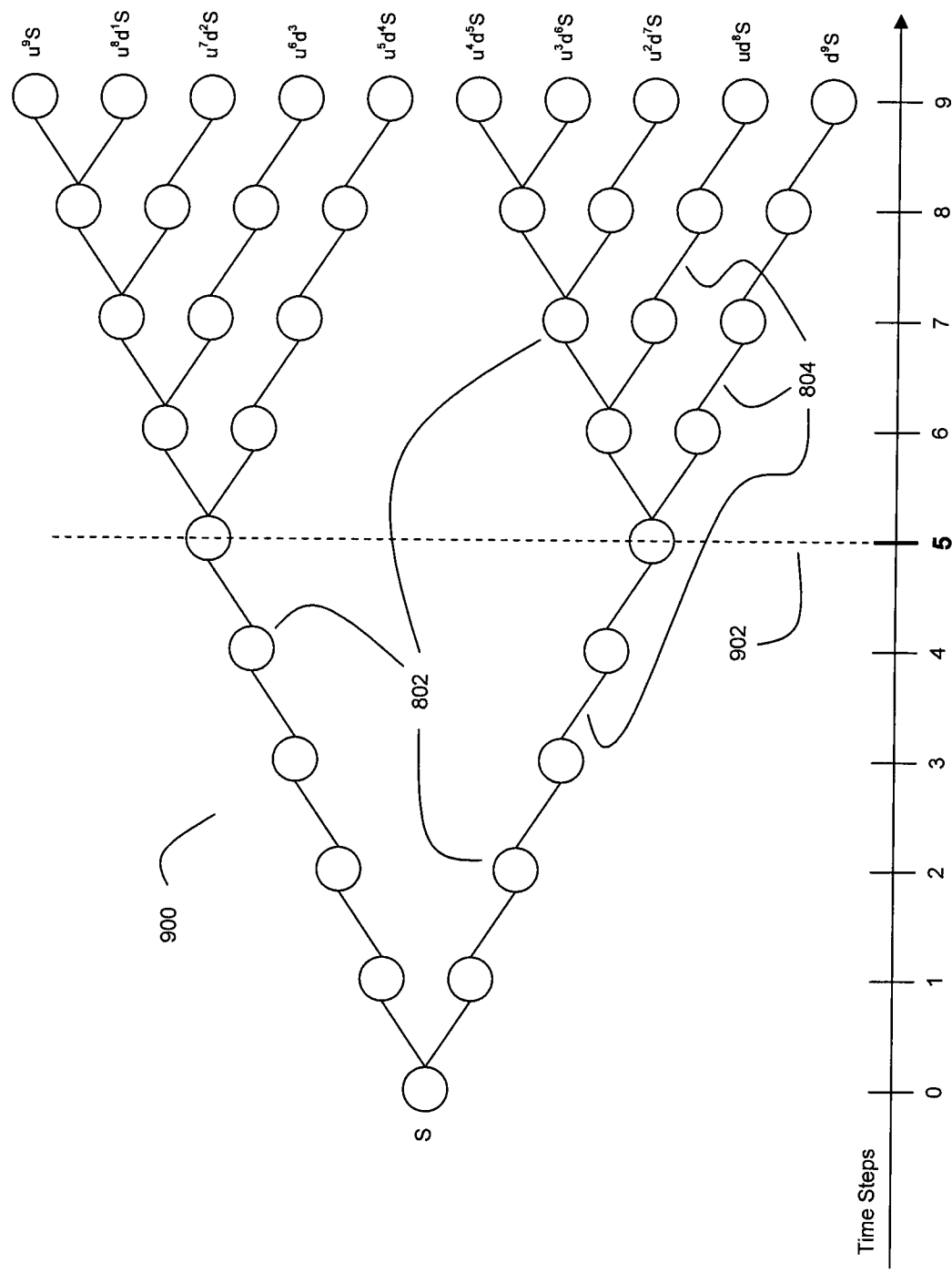


Figure 9(a)

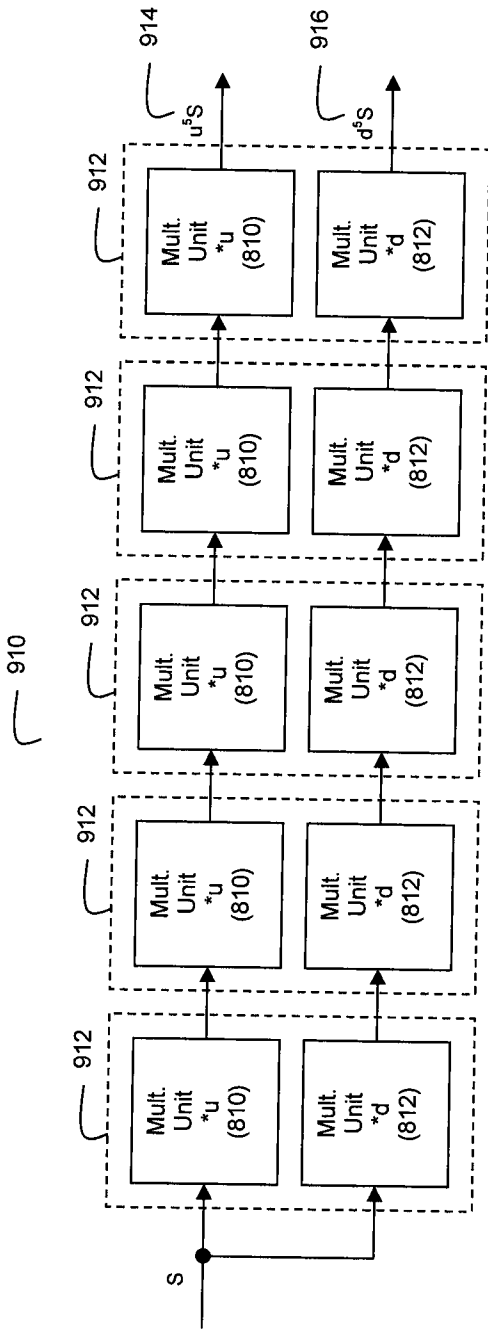


Figure 9(b)

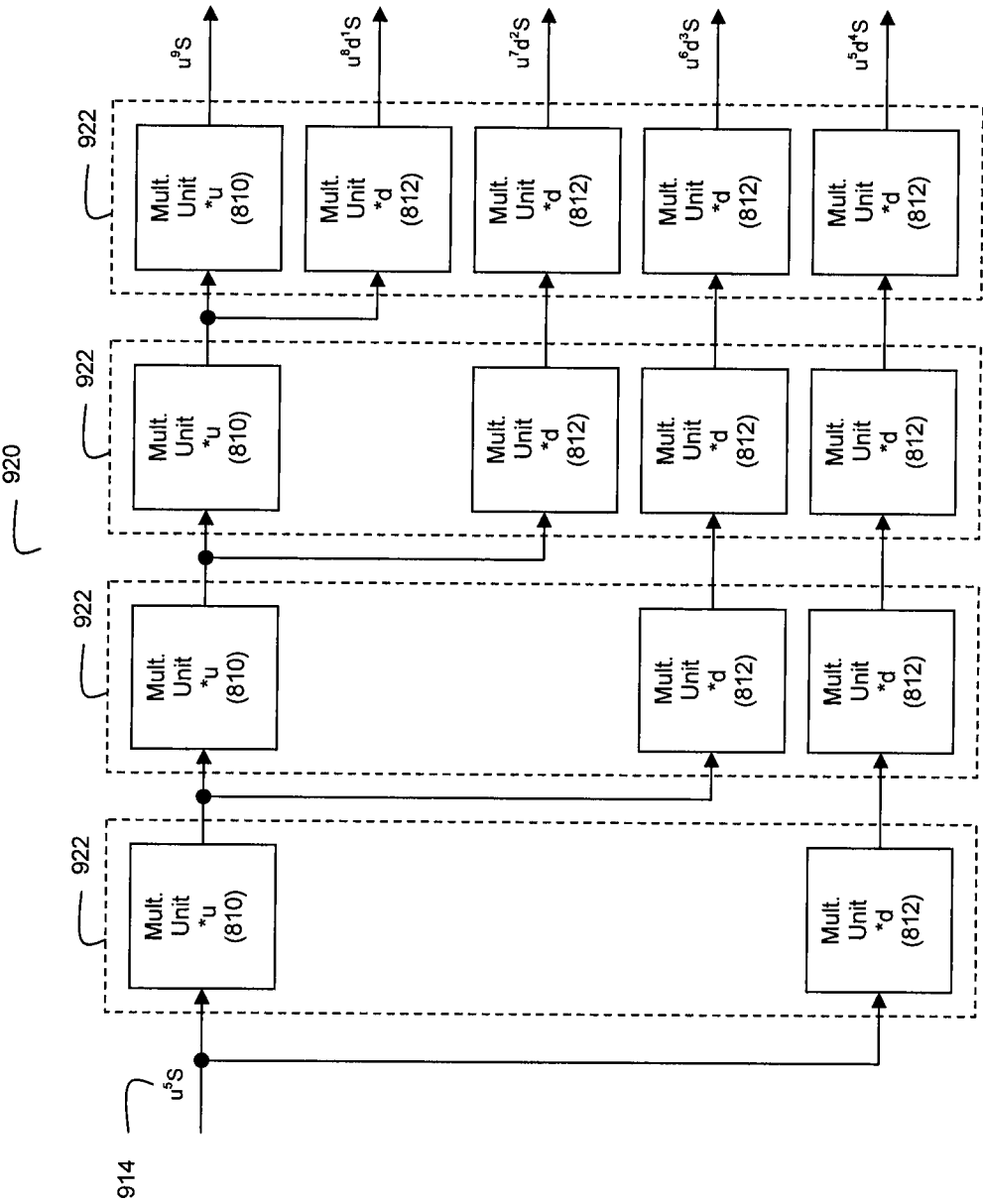


Figure 9(c)

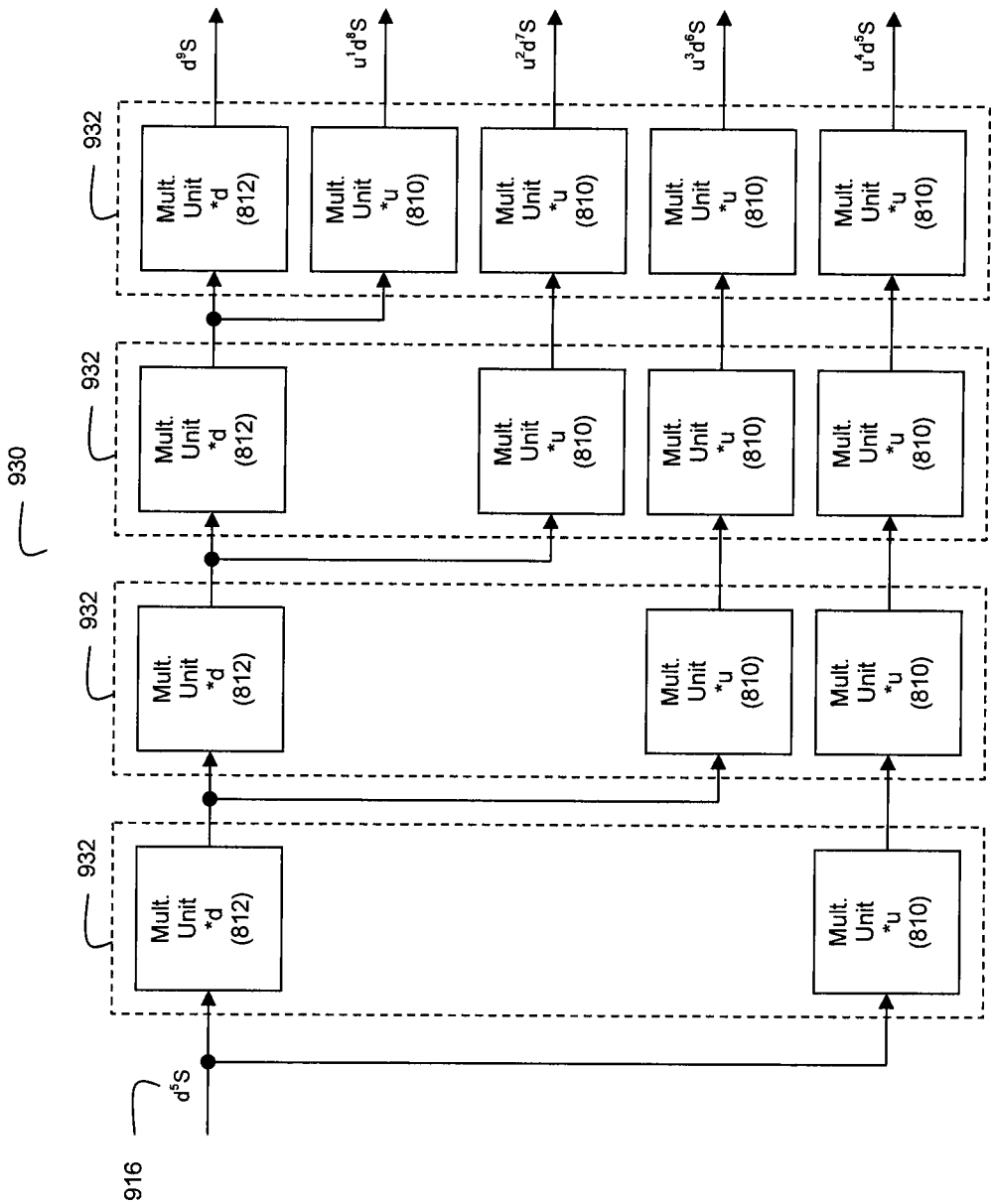


Figure 9(d)

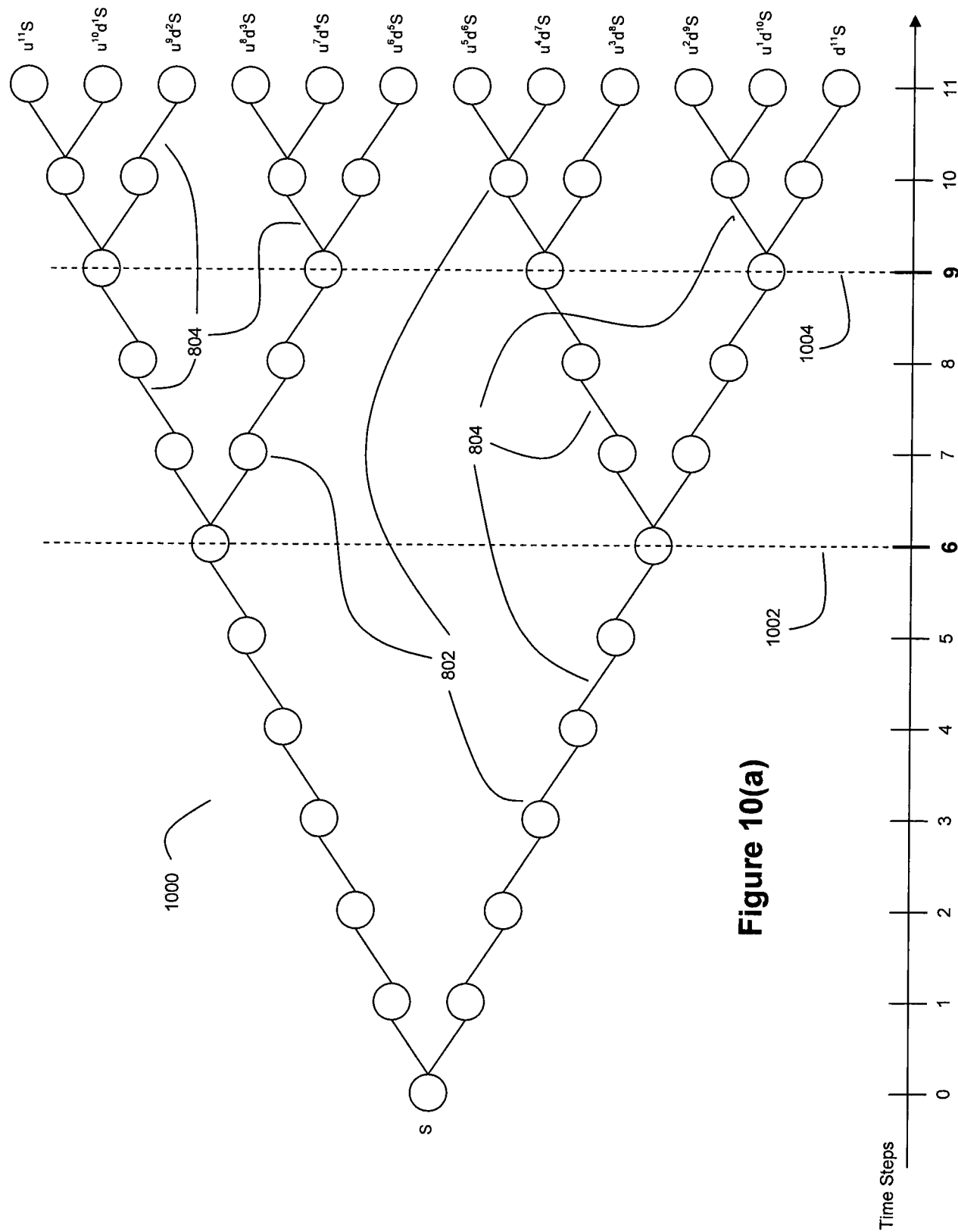


Figure 10(a)

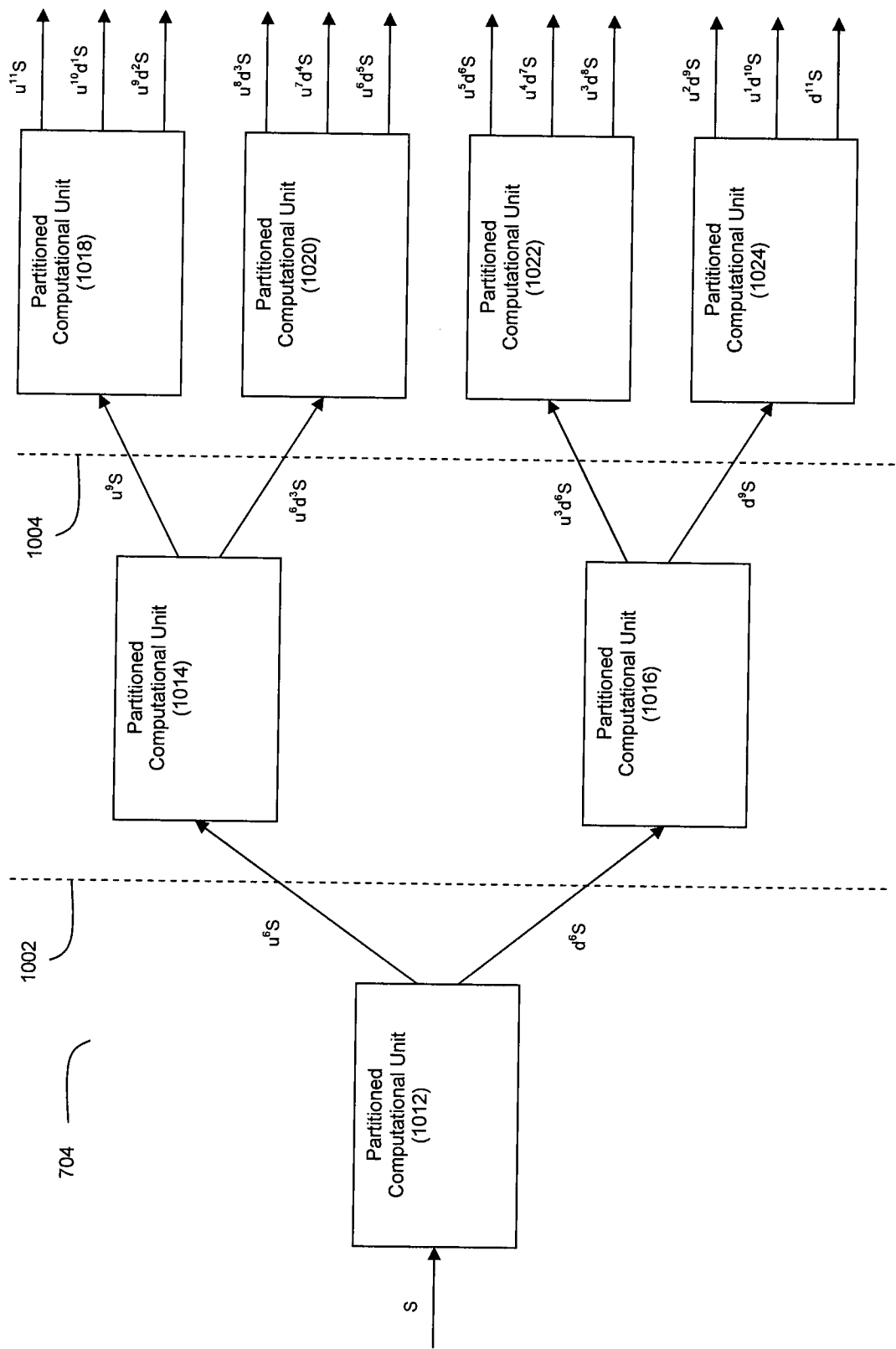


Figure 10(b)

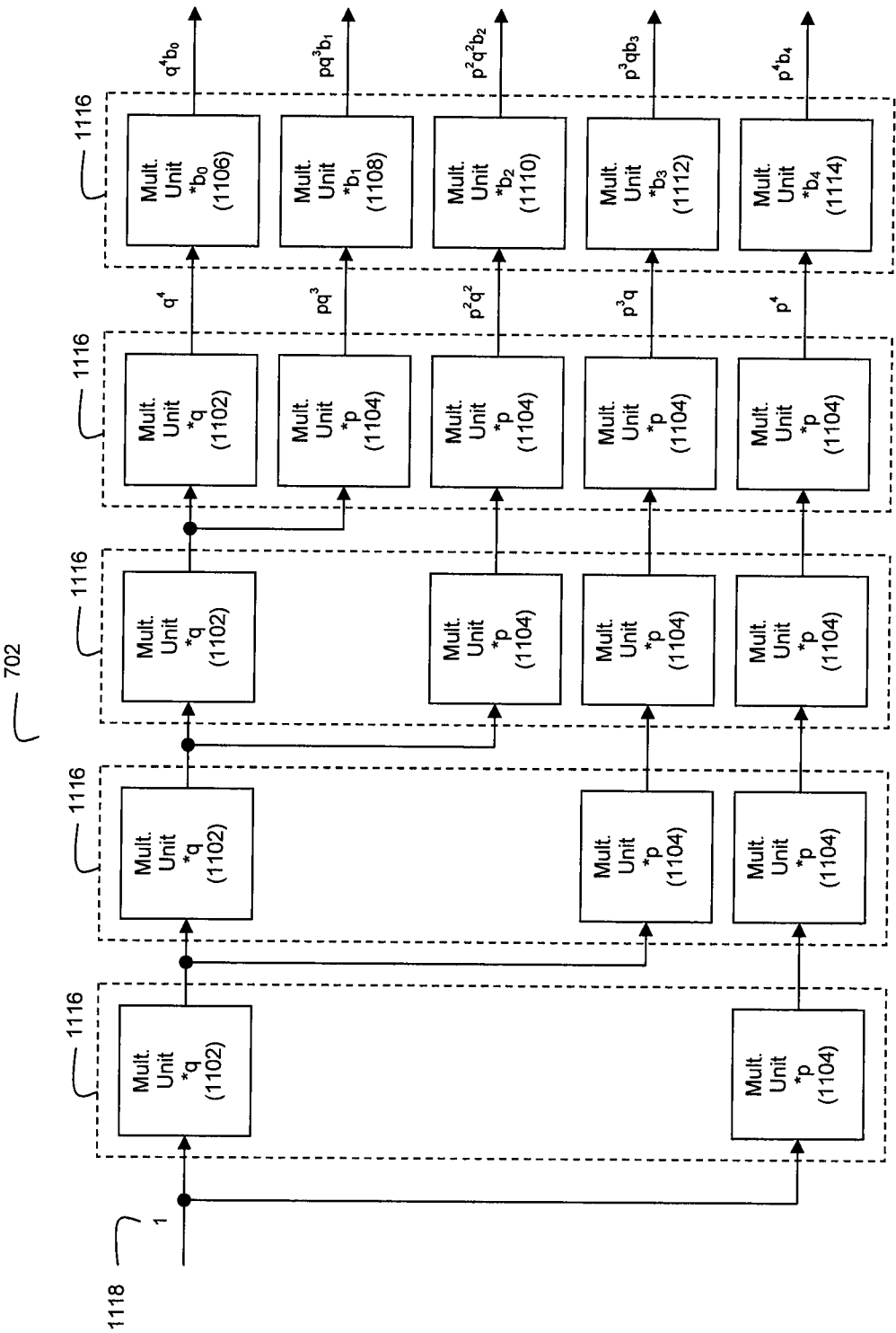


Figure 11

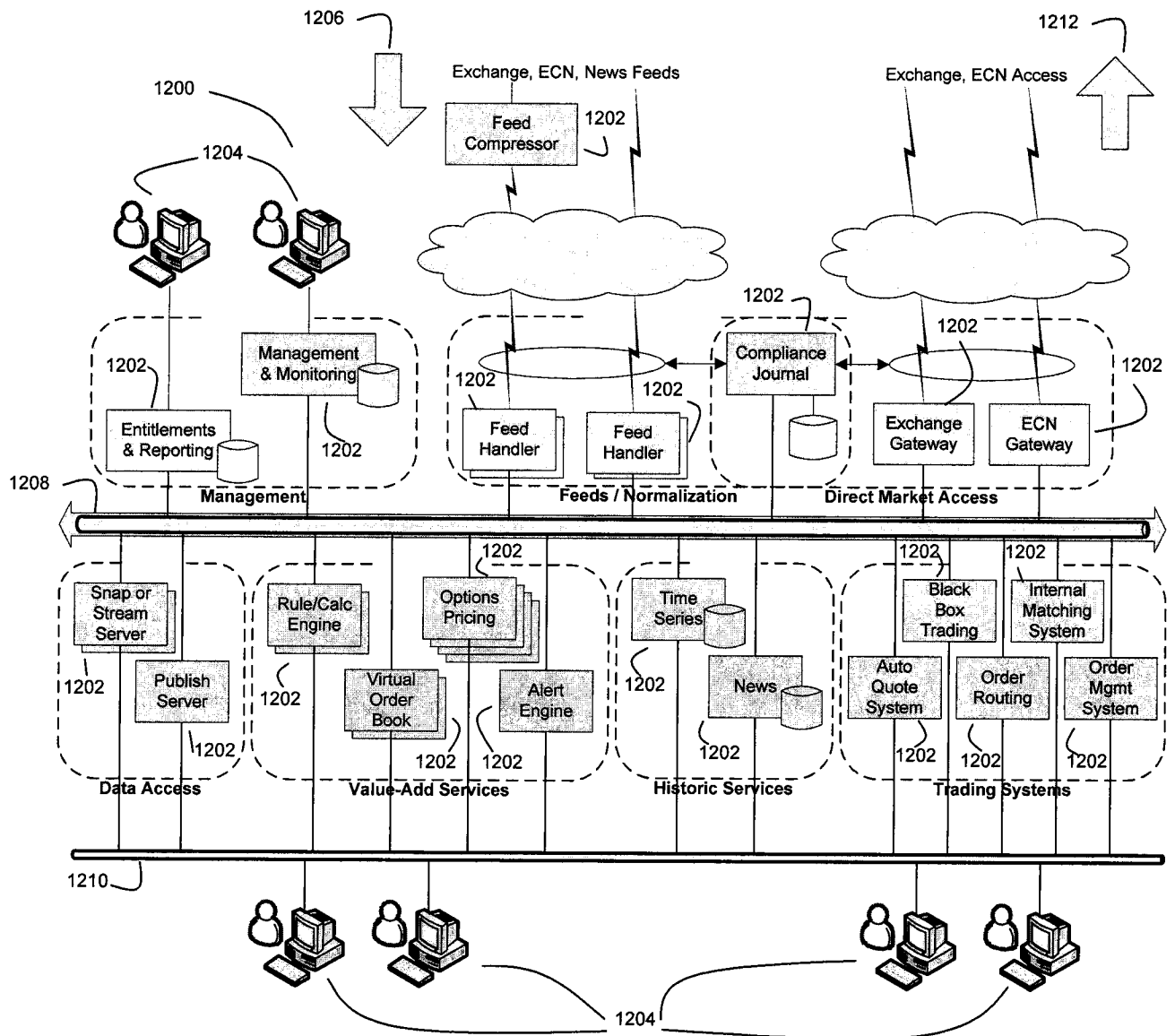


**Figure 12**  
**PRIOR ART**

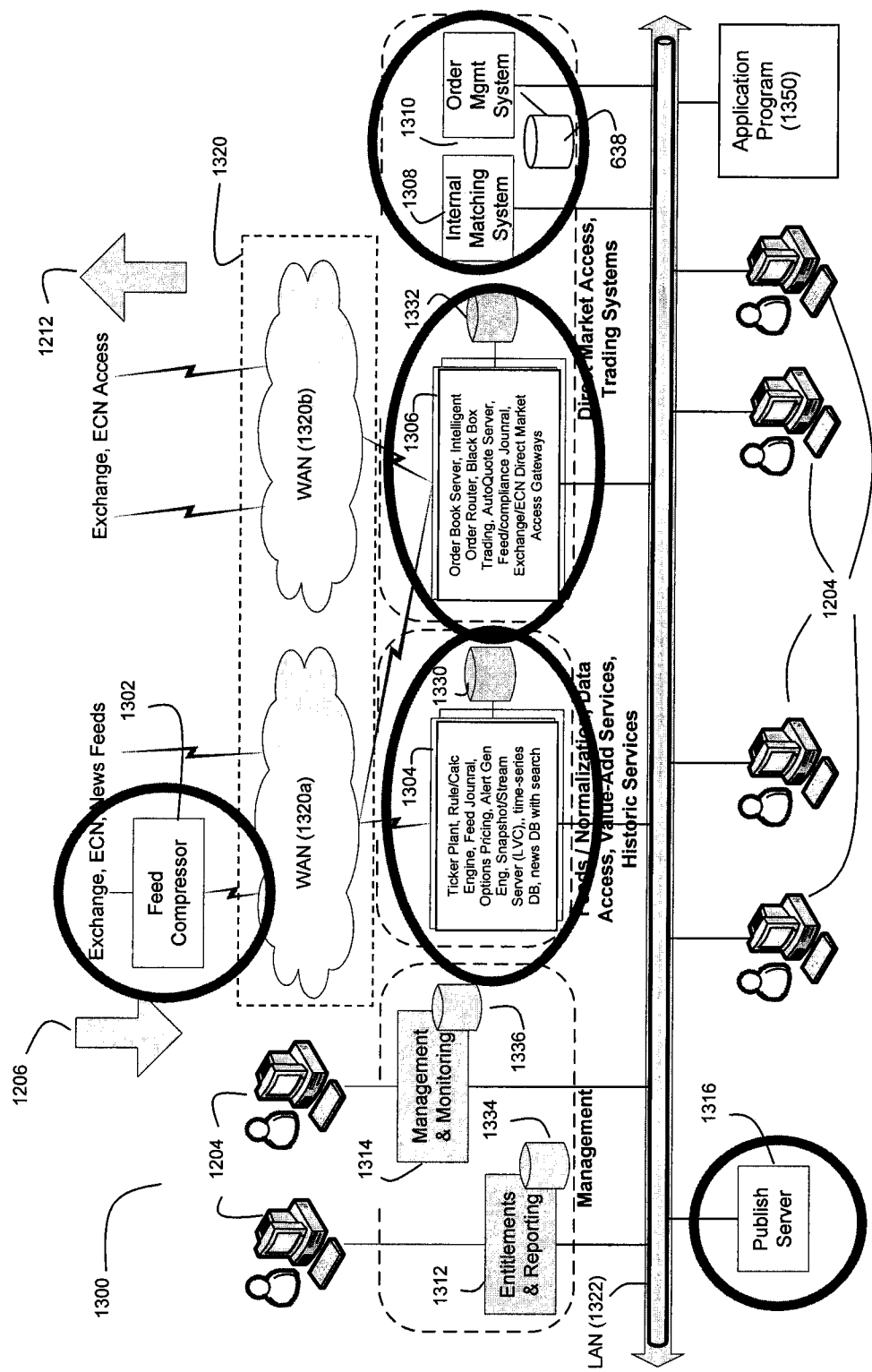


Figure 13

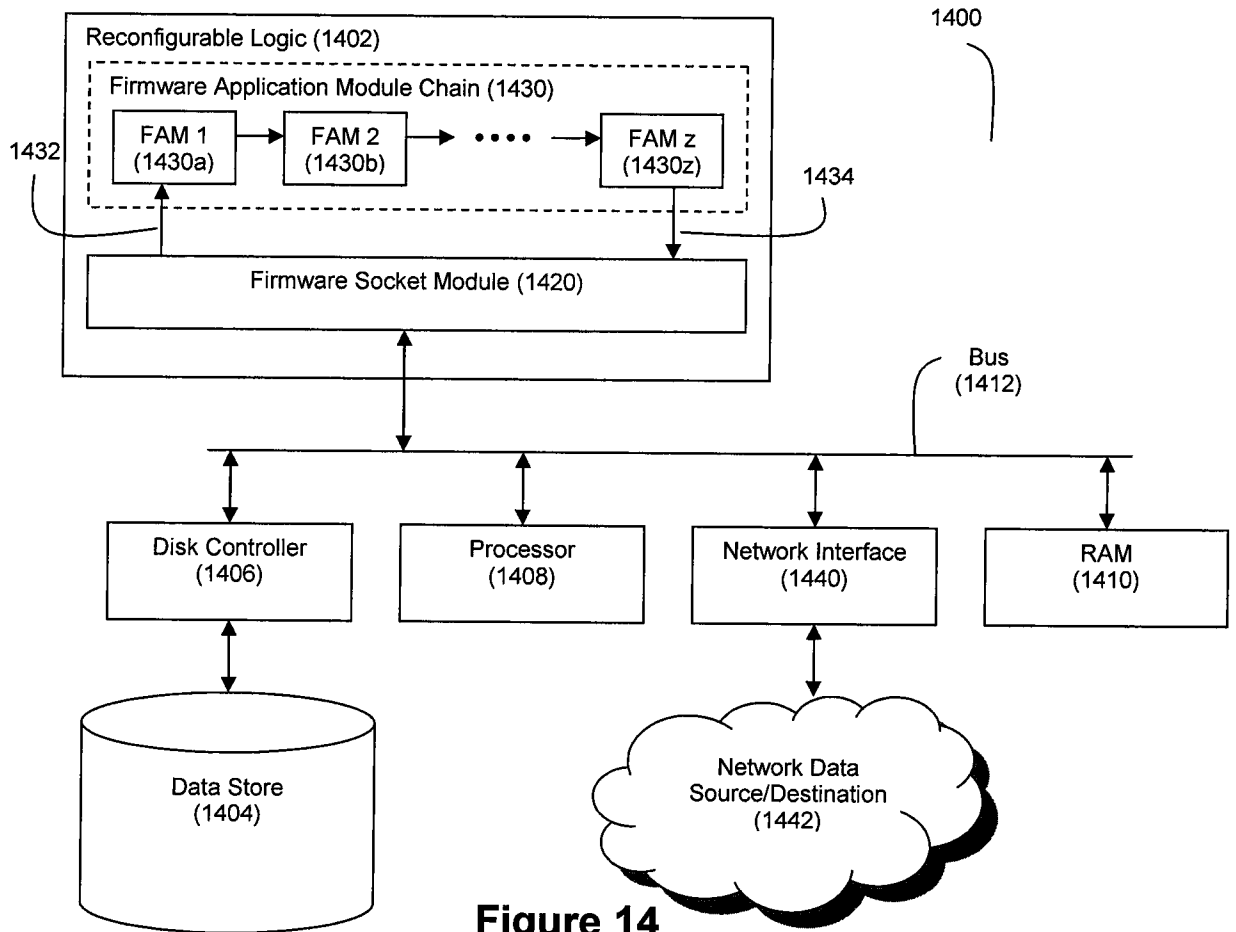


Figure 14

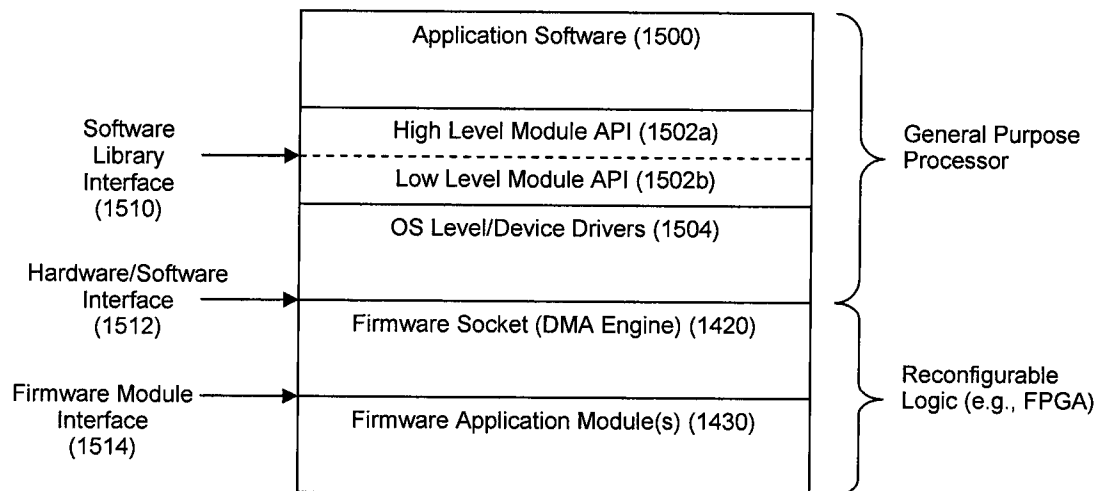


Figure 15

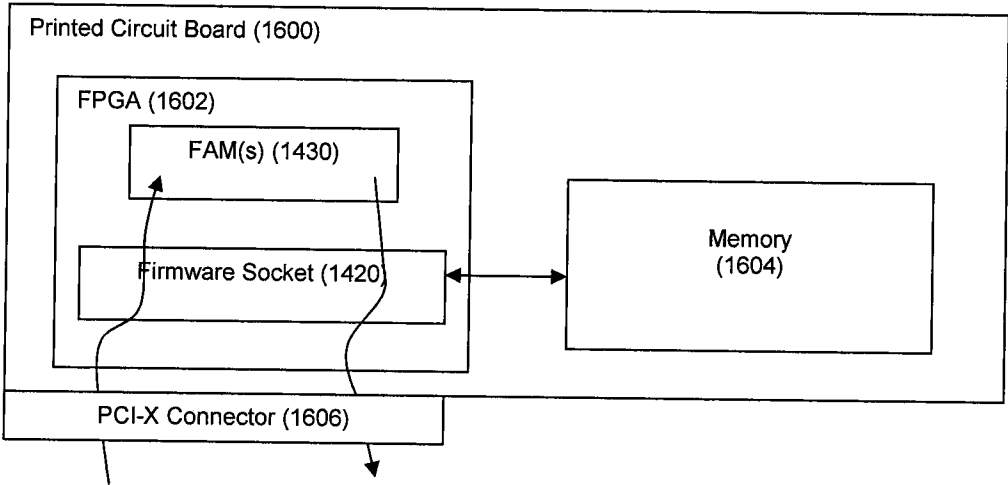


Figure 16(a)

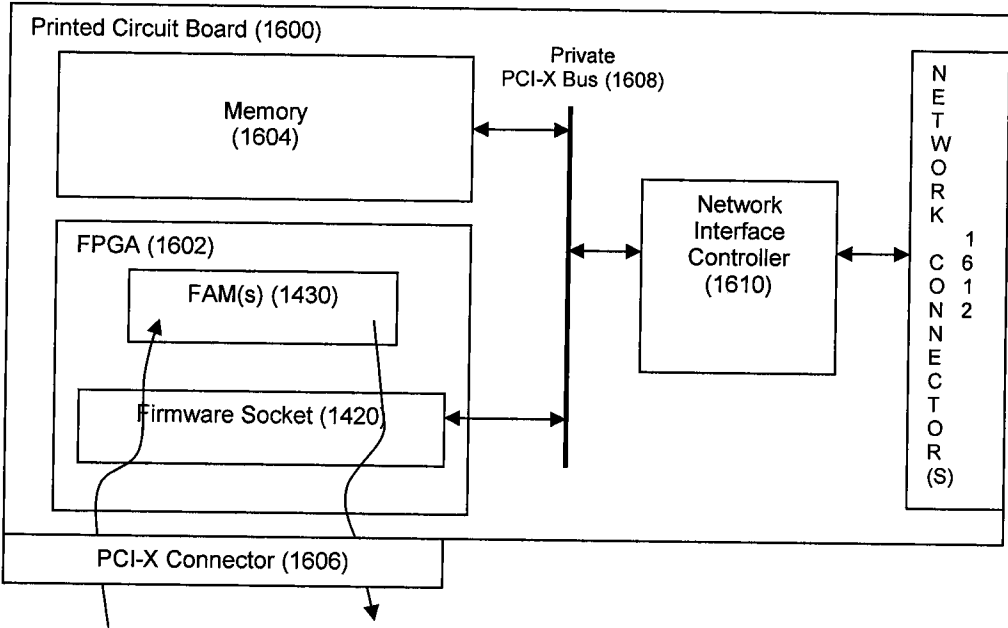


Figure 16(b)

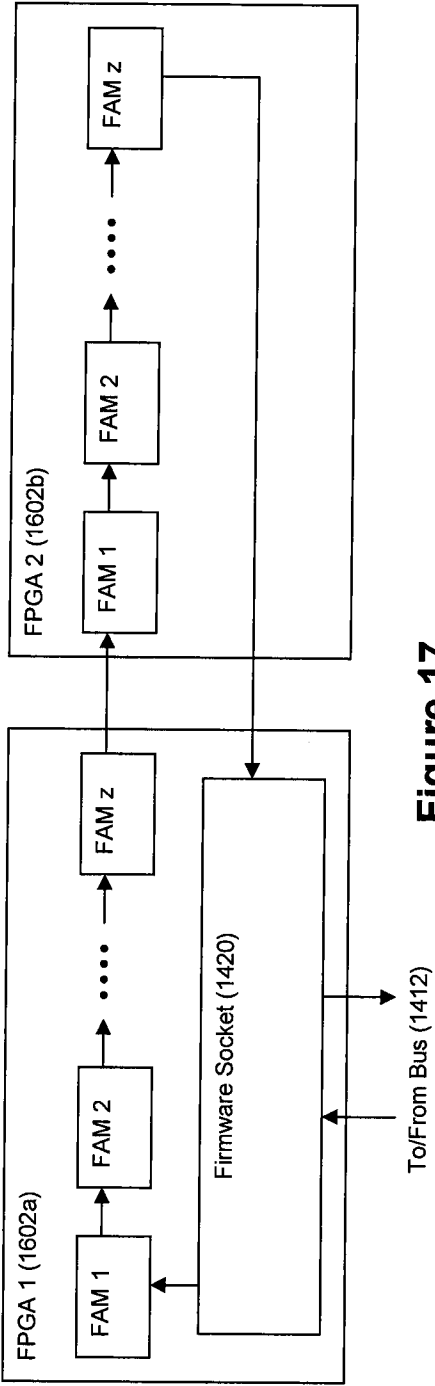


Figure 17

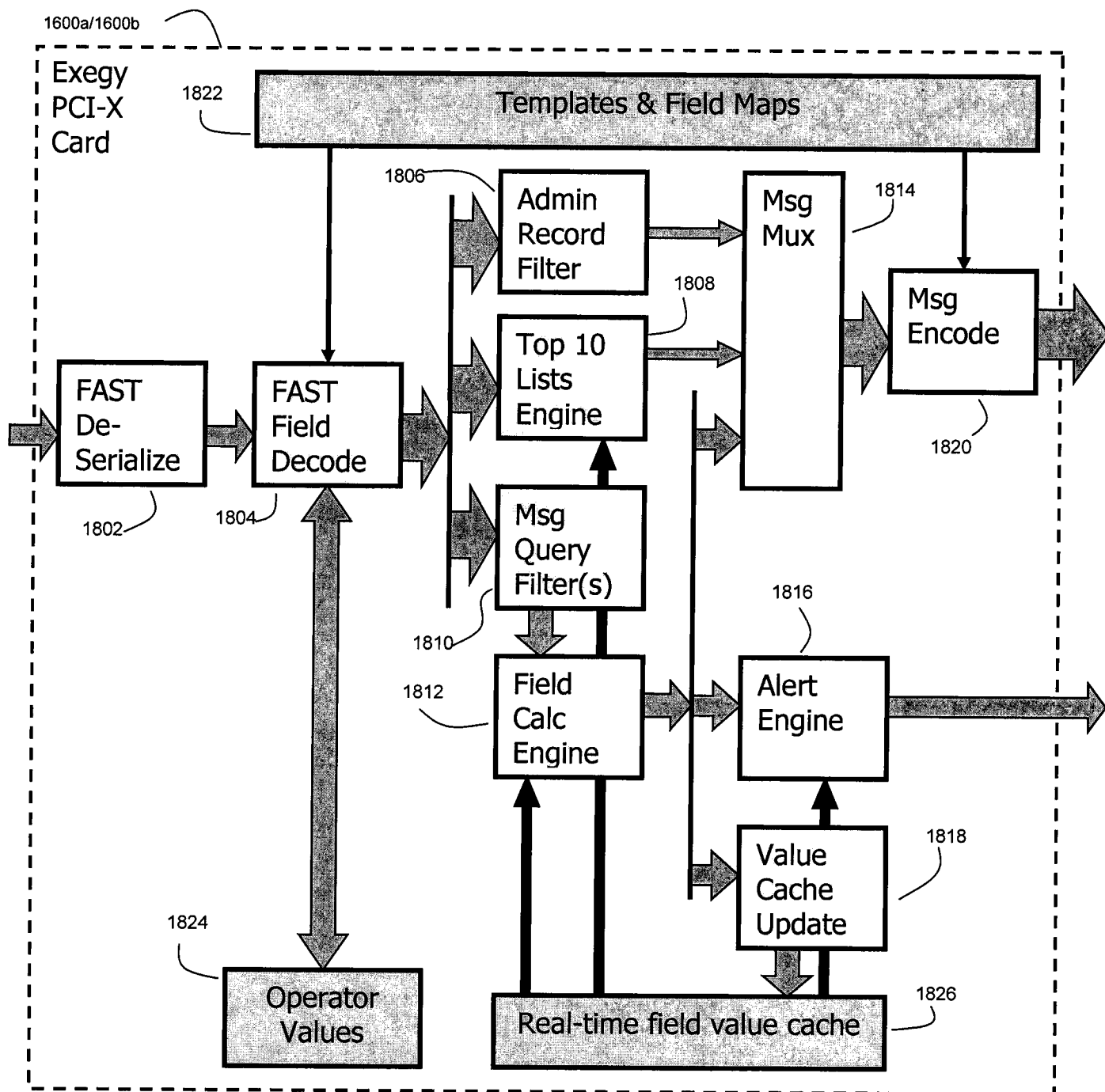


Figure 18

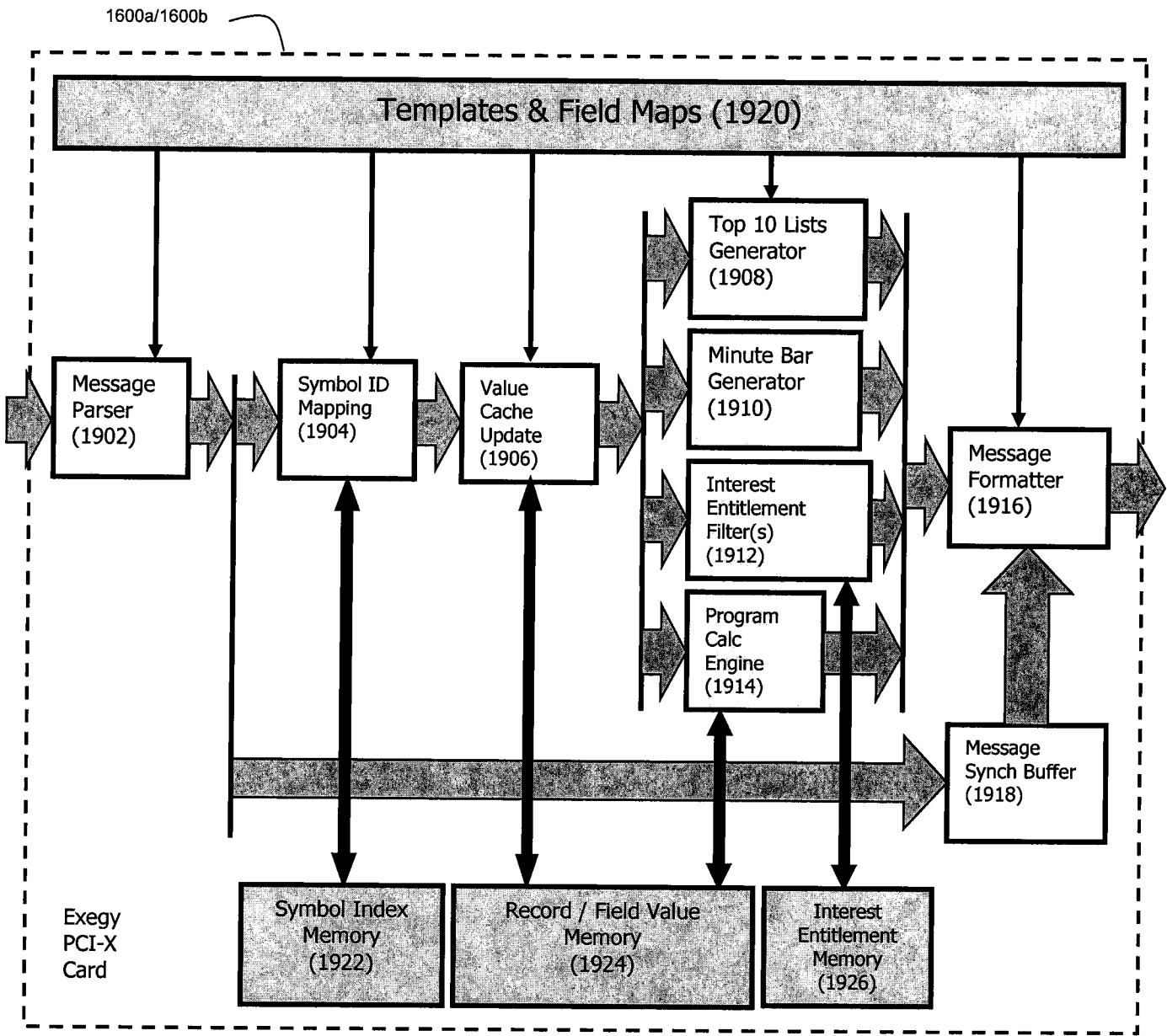


Figure 19

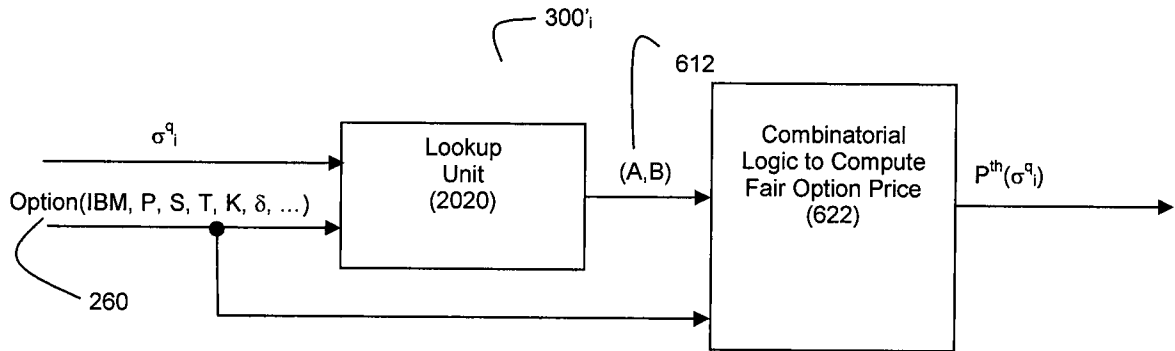


Figure 20(a)

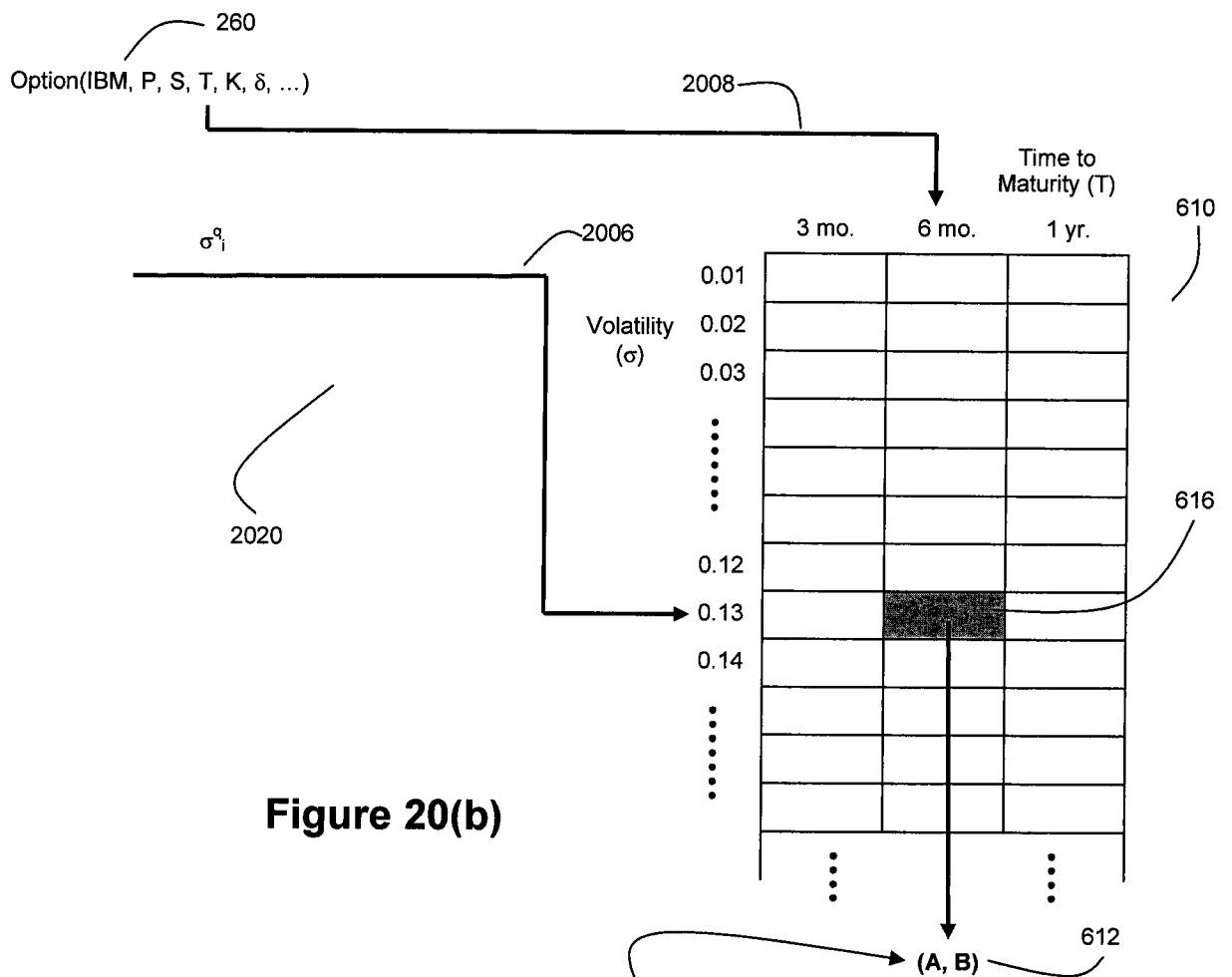


Figure 20(b)

The retrieved values for A and B are used in the computation of  $P^{th}(S, K)$ :

$$P^{th}(S, K) = \sum_j A_j \max(SB_j - K, 0)$$

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 08/65955

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - G06Q 40/00 (2008.04)

USPC - 705/36R

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8): G06Q 40/00 (2008.04)

USPC: 705/36R

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

USPC: 705/1, 35, 37; 700/1, 90

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Electronic databases: USPTO WEST (PGPB, USPT, EPAB, JPAB); Google Scholar

Search Terms Used: determining or calculating or computing option prices or volatilities, implied or intra-day volatility, parallel or pipelined architecture or processing, pricing model or CRR model, iterative or iteration execution or processing et

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category*     | Citation of document, with indication, where appropriate, of the relevant passages                                                            | Relevant to claim No.                                                                             |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|
| X<br>---<br>Y | US 2004/0034587 A1 (Amberson et al.) 19 February 2004 (19.02.2004)<br>(abstract, and para [0001]-[0009], [0019]-[0030], [0036]-[0044])        | 1-5, 7, 14-20 and 77-78<br>-----<br>27-29, 39-41, 54-60, 69-70, 79-84, 88-92, 99-103, and 112-119 |
| Y             | US 2003/0208430 A1 (Gershon) 06 November 2003 (06.11.2003)<br>(abstract, and para [0030]-[0040], [0045]-[0050], [0053]-[0057], [0065]-[0072]) | 27-29, 39-41, 54-60, 69-70, 79-84, 88-92, 99-103, and 112-119                                     |
| A             | US 7,149,715 B2 (Browne et al.) 12 December 2006 (12.12.2006)                                                                                 | 1-5, 7, 14-20, 27-29, 39-41, 54-60, 69-70, 77-84, 88-92, 99-103, and 112-119                      |
| A             | US 7,065,475 B1 (Brundobler) 20 June 2006 (20.06.2006)                                                                                        | 1-5, 7, 14-20, 27-29, 39-41, 54-60, 69-70, 77-84, 88-92, 99-103, and 112-119                      |

☐ Further documents are listed in the continuation of Box C.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

19 August 2008 (19.08.2008)

Date of mailing of the international search report

22 AUG 2008

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents  
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 08/65955

## Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☒ Claims Nos.: 6, 8-13, 21-26, 30-38, 42-53, 61-68, 71-76, 85-87, 93-98, 104-111, and 120-122  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.