Cluster Computing for Financial Engineering

Thomas F. Coleman

Professor, Computer Science & Applied Mathematics

Director, Cornell Theory Center

December 3, 2003.

Cornell Theory Center

Mission:

Advance and facilitate the use of industry-standard cluster technology to solve computational intensive problems in science, engineering, finance and business.

A center of excellence in high-performance computing and interdisciplinary research at Cornell University

- Over 1600 processors
(Windows clusters)
- World’s fastest..
- World’s largest..

Cornell Theory Center

Mission:

Advance and facilitate the use of industry-standard cluster technology to **solve computational intensive problems in science, engineering, finance and business.**

A center of excellence in high-performance computing and interdisciplinary research at Cornell University

- Over 1600 processors (Windows clusters)
- World's fastest..
- World's largest..

CTC Manhattan

Our Mission:

Promote technology transfer to industry.

1. **A**dvance industry-standard cluster computing solutions for our industry clients.

2. **D**evelop effective solutions for the problems of computational finance for our clients (often, cluster solutions).

My impressions of the use of parallel computing in this industry…

My impressions of the use of parallel computing in this industry…

- The big investment banks on Wall Street do use parallel computing but in a limited and distant way, usually remote from either active portfolio management or front-line analysis. Often batch

My impressions of the use of parallel computing in this industry…

- The big investment banks on Wall Street do use parallel computing but in a limited and distant way, usually remote from either active portfolio management or front-line analysis. Usually batch
- The use of parallelism declines rapidly as you move away from Wall street and away from the top investment banks (but the needs do not).

My impressions of the use of parallel computing in this industry…

- The big investment banks on Wall Street do use parallel computing but in a limited and distant way, usually remote from either active portfolio management or front-line analysis. Usually batch
- The use of parallelism declines rapidly as you move away from Wall street and away from the top investment banks (but the needs do not).
- Financial institutions ‘cheat’ – e.g.,often reducing the # of MC simulations needed (dramatically) to get the required ‘number’ by 9AM

My impressions of the use of parallel computing in this industry…

- The big investment banks on Wall Street do use parallel computing but in a limited and distant way, usually remote from either active portfolio management or front-line analysis. Usually batch
- The use of parallelism declines rapidly as you move away from Wall street and away from the top investment banks (but the needs do not).
- Financial institutions ‘cheat’ – e.g.,often reducing the # of MC simulations needed (dramatically) to get the required ‘number’ by 9AM
- Even small players could well benefit from a convenient parallel environment – **convenience and ‘desktop access’ far more important than ‘full’ parallel capabilities.**

..

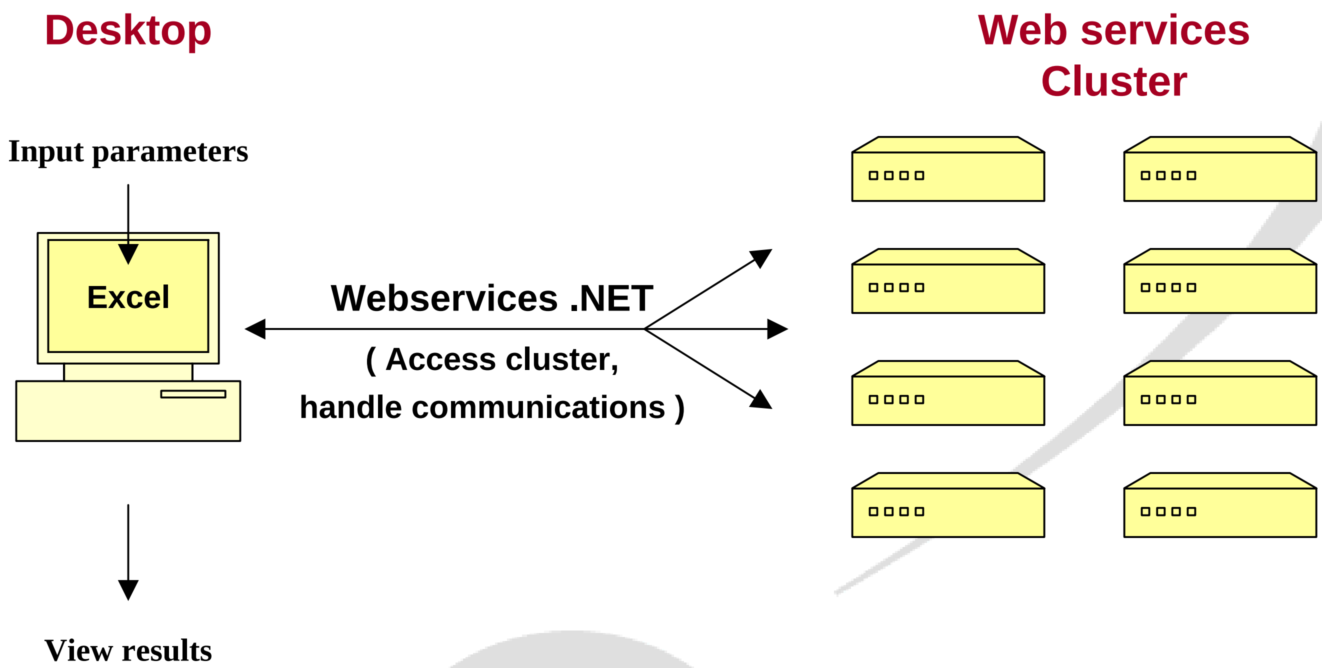
My impressions of the use of parallel computing in this industry…

- The big investment banks on Wall Street do use parallel computing but in a limited and distant way, usually remote from either active portfolio management or front-line analysis. Usually batch
- The use of parallelism declines rapidly as you move away from Wall street and away from the top investment banks (but the needs do not).
- Financial institutions ‘cheat’ – e.g.,often reducing the # of MC simulations needed (dramatically) to get the required ‘number’ by 9AM
- Even small players could well benefit from a convenient parallel environment – convenience and ‘desktop access’ far more important than ‘full’ parallel capabilities.
- Growing interest in ‘grid computing’ but in a very limited way..

Our Quest:

Efficiently solve compute-intensive practical problems in computational finance, using parallelism, under web services.

The Target Environment:



Strengths and Weaknesses of this View

Strengths:

- Driven by application/user
- Relatively high availability environment
- ‘extra’ communication software/machinery is not needed.
- Rich environment – commercial, tools, packages etc..

↓

Immediate industry interest

Strengths and Weaknesses of this View

Strengths:

- Driven by application/user
- Relatively high availability environment
- ‘extra’ communication software/machinery is not needed.

Weaknesses (challenges)

- **Master-worker parallelism only / limited communication**
- **Saving state?**
- **Job scheduling?**

Snapshot of Computational Finance (what do 'quants' do?)

What are the computing problems of finance?

1. Computing the 'fair price' of financial instruments (based on 'no arbitrage' principles.) Implied vol, vol surfaces. **Future values.**
2. **Computing sensitivities, risk factors**
3. **Hedging**
4. Portfolio construction, rebalancing
5. Portfolio valuation, sensitivity
6. Index tracking
7. **Value at risk, Conditional value-at-risk**

The numerical math:

- *a) PDEs*
- *b) Optimization (continuous, stochastic). LPs,QPs, nonlinear,..*
- *c) Monte carlo simulations*
- *d) Regression, fitting, inverse problems*
- *e) Splines, approximations*

A FE min-Intro….

Movement of stocks

Various models to capture movement of stocks

A Basic and popular model: the Geometric Brownian Motion model:

$$\boxed{\frac{dS_t}{S_t} = \mu \cdot dt + \sigma \cdot dW_t}$$

W: Brownian motion

μ: constant, the drift

σ: constant, the volatility

What is an option?

An option is a contract based on an underlying (e.g., stock) with a choice to exercise or not.

Example, European **call option**: Today you agree to **have the option** to buy a stated stock at future time **T** (expiration time) at strike price **K**.

Option examples

European **put option**: Today you agree to have **the option** of selling a stated stock at future time **T** (expiration time) at strike price **K**

American **call option**: Today you agree to have **the option** to buy a stated stock at **or before** a future time T (expiration time) at strike price **K**

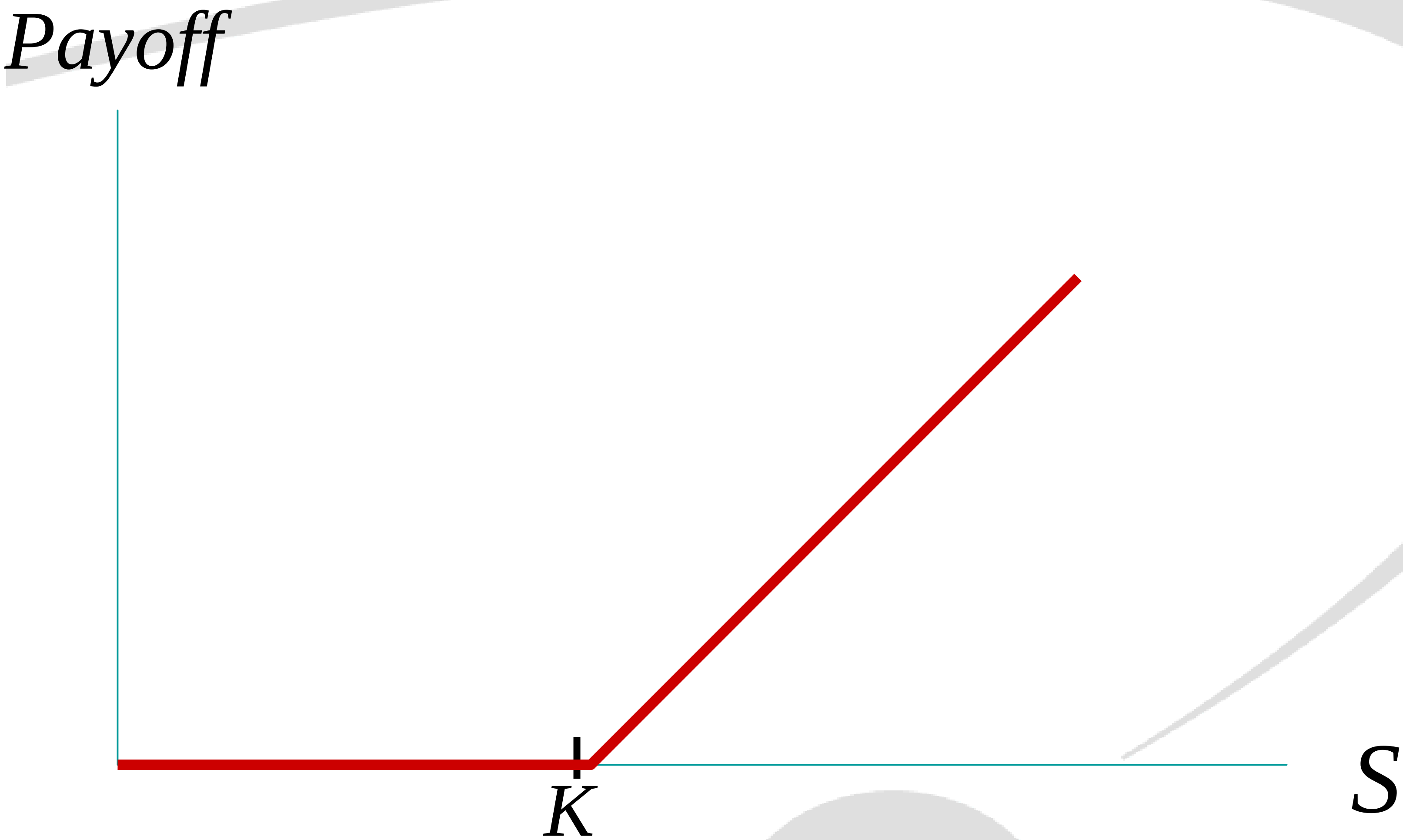
More option examples

Asian option: options based on average asset price over some fixed period (path dependent)

Lookback option : based on maximum or minimum asset price over some period

Barrier options: option expires if (unless) asset crosses prefixed barriers (up-and-in, down-and-out,…)

European call payoff function



Central Questions about Options

What is a fair price of a prescribed option?

How do you effectively hedge with options?

Black-Scholes

Under a number of 'pure market' assumptions,including no arbitrage,GBM model, and complete market, the unique fair price of a vanilla European option satisfies the famous Black-Scholes equation:

Black-Scholes

Black-Scholes solution

$$\frac{\partial v}{\partial t} + (r - q)s\frac{\partial v}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 v}{\partial s^2} = rv$$

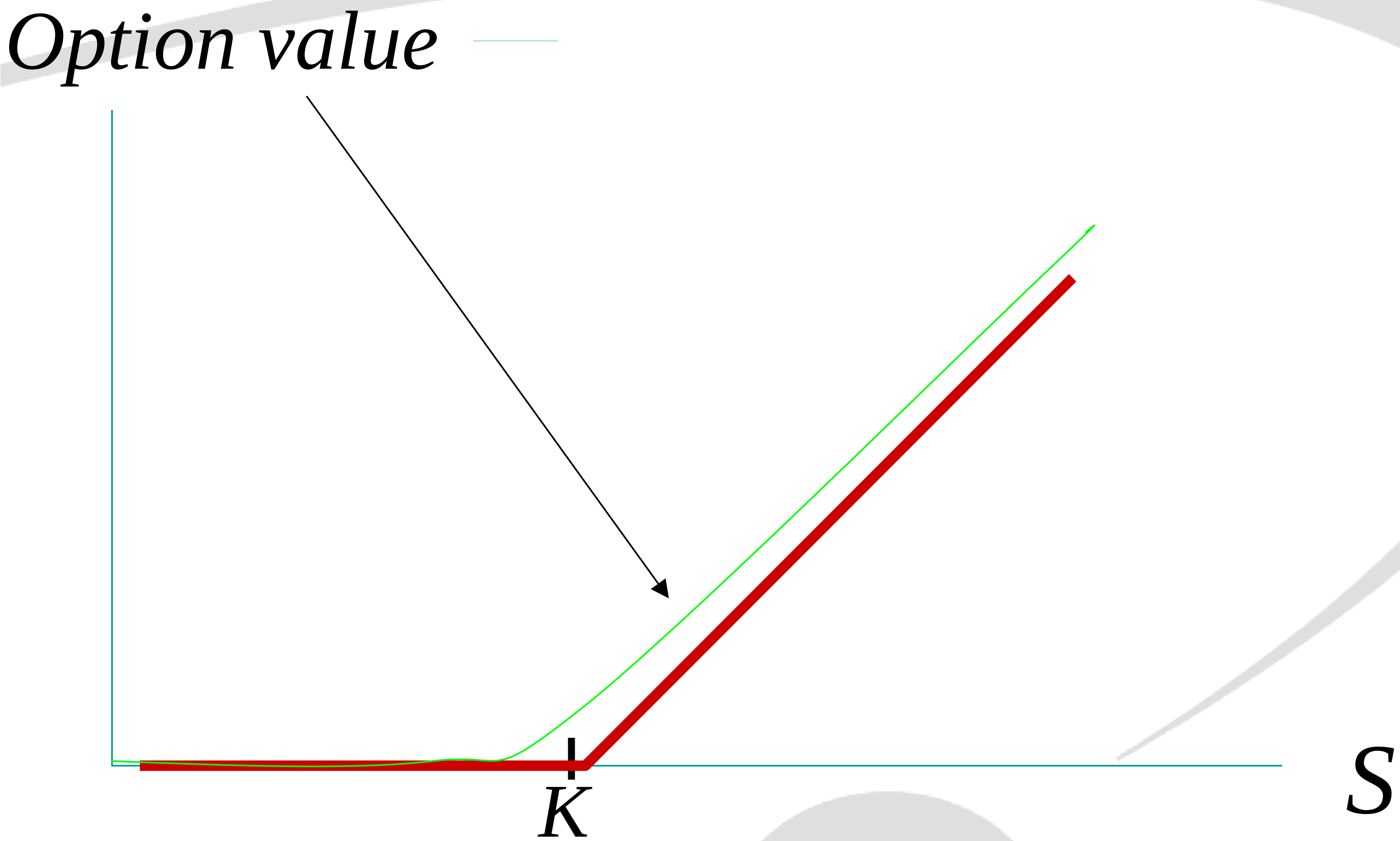
r: risk-free interest rate

q: dividend rate

+ natural boundary conditions

There exits explict analytic soln to this PDE…instantaneous.

European call option value



Implied Volatility

Inverse problem: Given price of an option what volatility value (scalar) allows B-S model to yield price:

- $$F(\sigma) - value = 0$$

PROBLEM: smiles and smirks!

Volatility Surface

But this leads to a non-constant (i.e., different data points yield diff. answers)

In fact, it appears

$$\sigma = \sigma(S_t, t)$$

Computationally, things now get more interesting...

Option pricing model: 1-Factor Continuous Diffusion Approach

$$\frac{dS_t}{S_t} = \mu(S_t, t)dt + \sigma\ (S_t, t)dW_t$$

W: standard Brownian motion

μ, σ: deterministic functions

$\sigma(s, t)$: local volatility function

Fair price for vanilla option :

Generalized Black - Scholes:

$$\frac{\partial v}{\partial t} + (r - q)s\frac{\partial v}{\partial s} + \frac{1}{2}\sigma(s,t)^2 s^2 \frac{\partial^2 v}{\partial s^2} = rv$$

Evaluation

Given the vol surface $\sigma = \sigma(S_t, t)$

Numerical approaches can be used to solve the generalized B-S equation. But,

How to get $\sigma = \sigma(S_t, t)$?

An optimization solution

The problem $$\min f(\bar{\sigma}) \equiv \left\| F(v(\bar{\sigma})) \right\|^2$$

Is nonlinear least-squares (as before). But the number of unknowns is the number of knot points *p* which can be chosen

$$\boxed{p \leq \text{ \#option values available}}$$

A bit of 'art' needed here

The message: modern practical mathematical models have moved away from simple analytic formulae – increasingly sophisticated computational approaches are required

Cluster Example 1: Single instrument Callable Bond

A Single-Instrument Example: Pricing a Callable Bond

A bond is callable if it can be 'called in' at any of a number of fixed set times over the lifetime of a bond (established at issuance).

The fair price of the bond depends on **unknown interest rates** in the future (over the lifetime of the bond), coupon payments (fixed), and **optimal 'call-in' policy** (which depends on the interest rate).

The callable feature makes this a 'American style', or, more precisely, Bermudan style, instrument

How to price? How to price faster (using parallelism)?

Pricing a callable bond

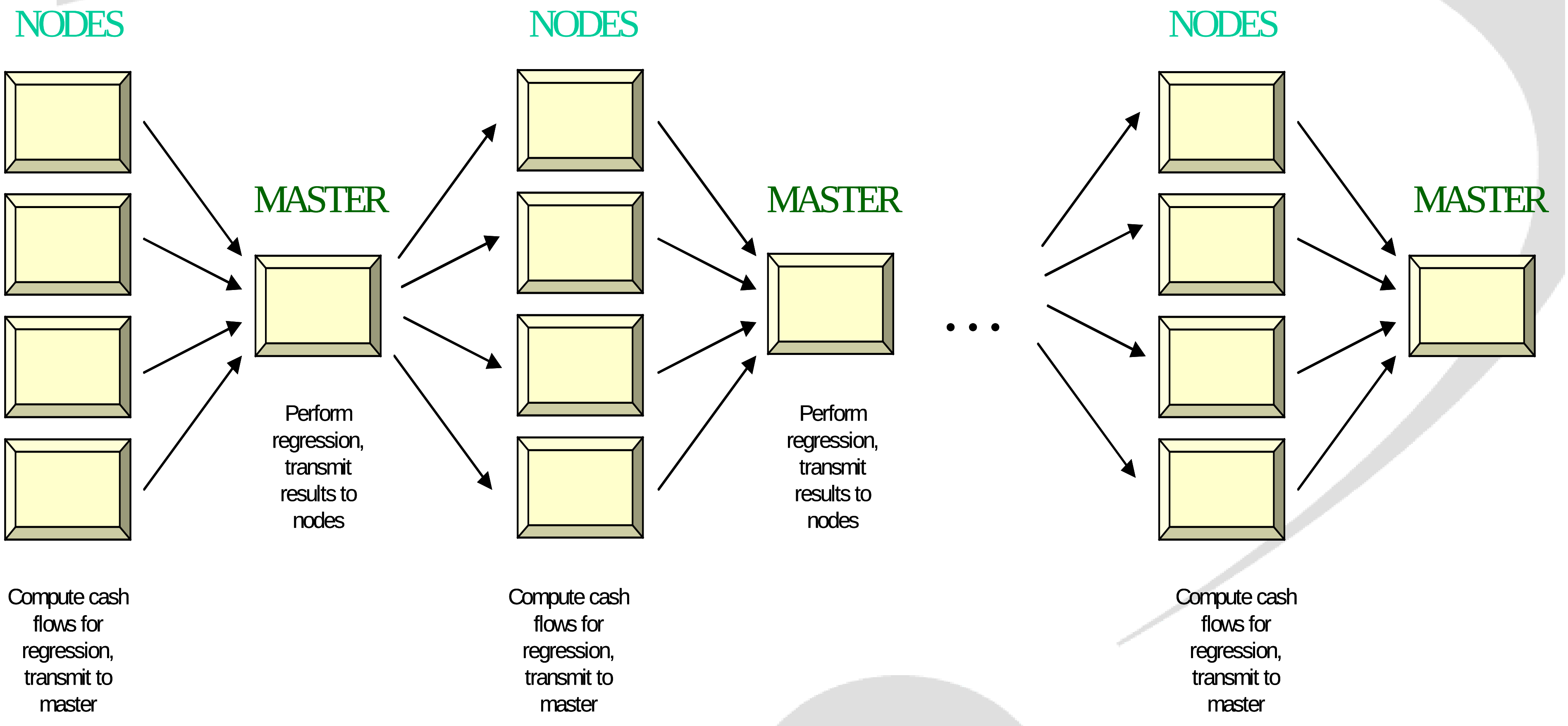
- **V**ery complex to price American style feature – exercise decision and current value depend on stochastic interest rates
- **L**east Squares Monte Carlo (LSM)
  - Longstaff and Schwartz (2001)
  - Equips Monte Carlo method with regression analysis to compute fair prices of any future cash flow.
  - Callable bond price = Non-callable bond price – call option price

The Algorithm (essence)

1. Forward pass: MC – generate interest rate (IR) scenarios, coupon prices

2. Backward pass: propagate values and solve a sequence of regression problems (at callable times).

Parallelize: Divide up the 20k IR scenarios. Limited communication is required for each regression. Master-worker.


NODES
MASTER
NODES
MASTER
NODES
MASTER
Perform regression, transmit results to nodes
Perform regression, transmit results to nodes
Compute cash flows for regression, transmit to master
Compute cash flows for regression, transmit to master
Compute cash flows for regression, transmit to master

Demo 1: Single Instrument LSM…

Results…

Speedup is quite limited due to communication/computation balance…

But many practical pricing problems are really at the portfolio level…

A Hierarchy of Complex Pricing/Evaluation Problems

Increasing Computational Cost

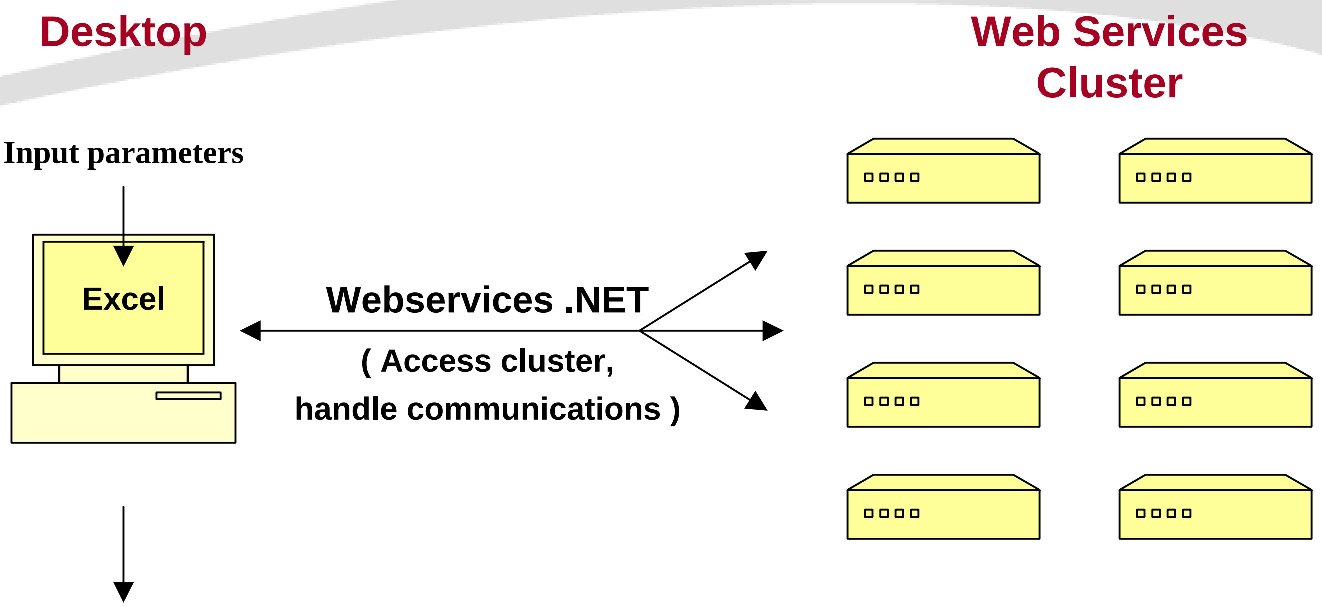
Find investment distribution in the portfolio that *minimizes* VaR

Compute VaR of the portfolio

Price a portfolio of complex bonds

Price a complex bond

Parallel computing on Windows-clusters driven from the desktop



View results

Monte Carlo interest rate simulations, company-specific simulations of default, and computation of future bond price distributions

Three PORTFOLIO demos:

1 Callable bonds
2 Risky bonds
3 Convertibles

Environment specifications

Compute nodes (running web services under Windows (.Net)):

- DELL Poweredge 2450's
- Dual Intel Pentium III processors
- 933 MHz
- 2. GB local RAM per node

Communication:

LAN (100Mbs) + internet (100Mbs [depends on load])

Master:

laptop: DELL latitude C400, 788 MHz, 512 MB of RAM

Note: no ownership of nodes…

Portfolio Example 1: Pricing Portfolios of Complex Bonds (on a Windows cluster, under .Net, using Web Services)

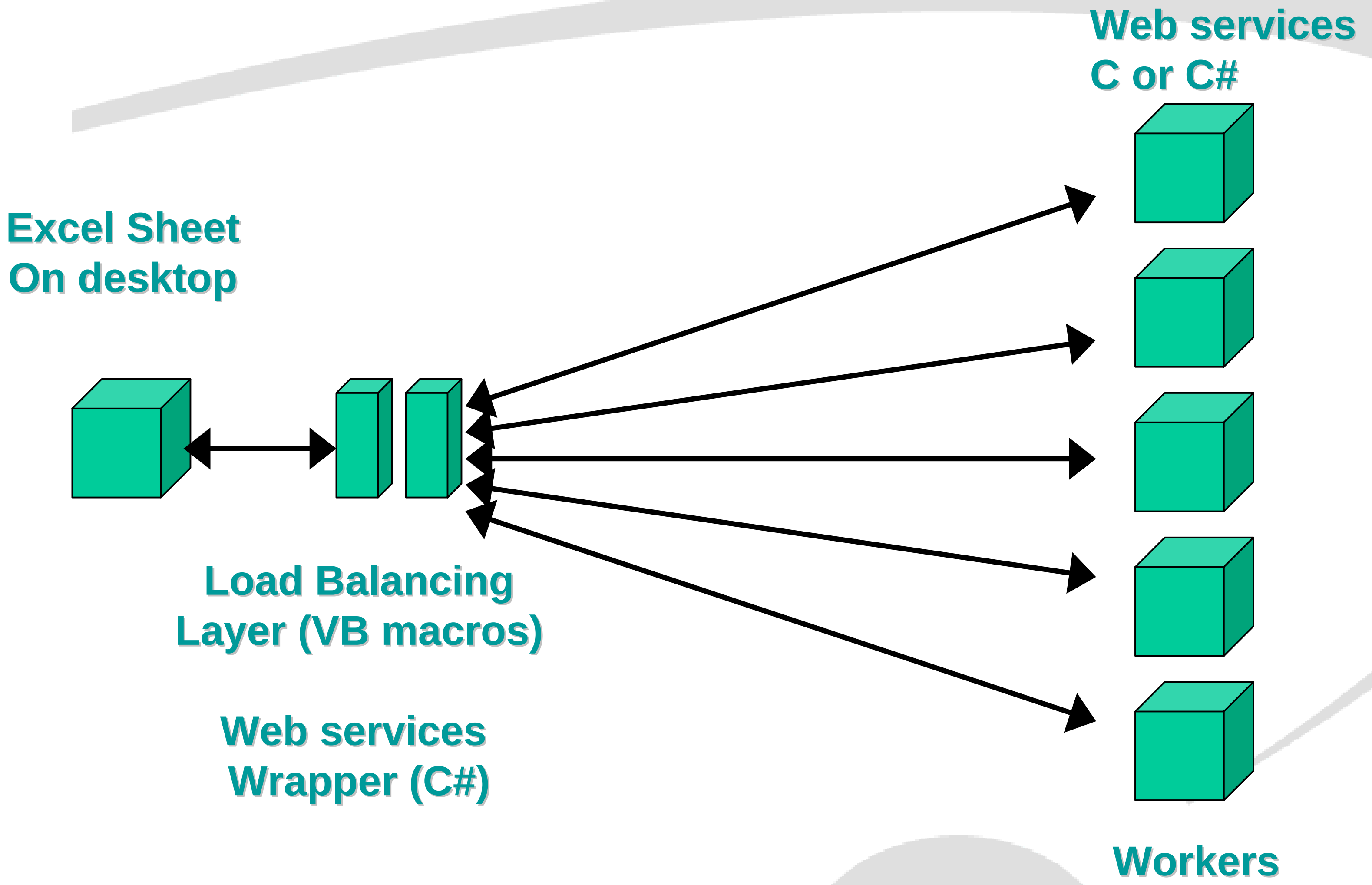
Key observation:

Some minor (occasional) redundancy in step 1 is ok…

- **Forward pass: MC – generate interest rate (IR) scenarios +**

2. **Backward pass: propagate values and solve a sequence of regression problems (at callable times).**

Master-Worker Model for Portfolio Pricing



Portfolio Demo 1: Pricing a portfolio of complex bonds on a cluster

December 3, 2003

Thomas F. Coleman

Not all portfolio problems present the same challenges (re Parallelism)

Previous demo: easy parallelism best – excellent performance, little communication. Some redundant computations in some circumstances.

Next portfolio problem: Previous approach (next instrument goes to next available processor) does not work → too much communication.

Portfolio Example 2: Pricing a portfolio of risky bonds on a cluster

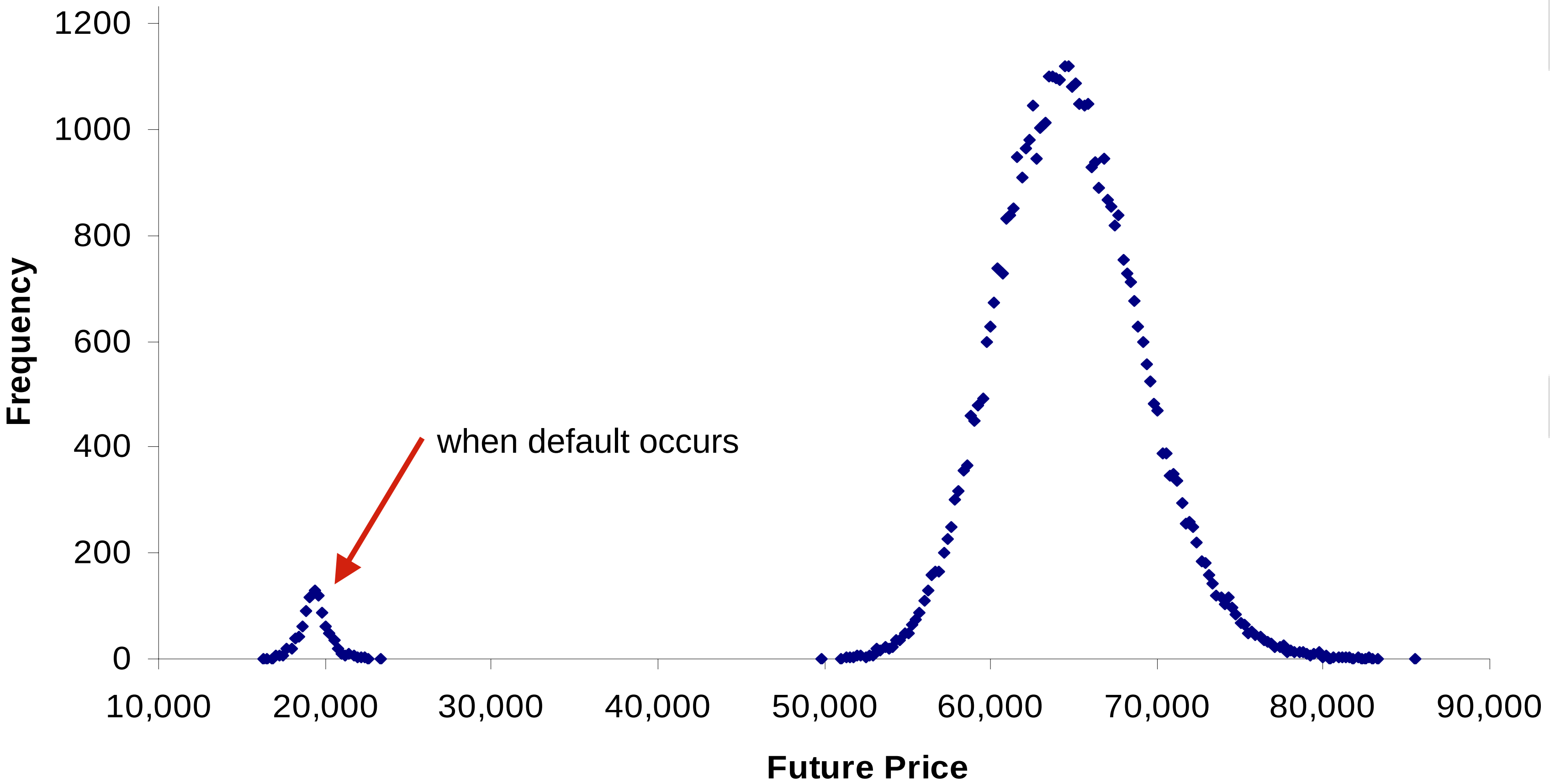
December 3, 2003

Thomas F. Coleman

Corporate Bonds

- **A**ttractive **–** higher yields than Treasury bonds
- **R**isky **–** credit risk: possibility the issuers will default – default possibility is a company specific measure – can be modelled/quantified through solving extensive regression problems
- **P**rice depends on default risk + (unknown) future interest rate → must simulate interest rate evolution and average to get a fair price
- **C**orporate bond managers – interested in fair price of portfolio + risk measures such as Value-at-risk (VaR)

Price Frequency Graph (with a bump)



The (parallel) cluster implementation

- **E**ach bond is priced (and Var, CVaR) calculated over entire cluster
- **M**onte-Carlo simulation of interest rate (Vasicek)
- **D**efault probability is measured (using historical corporate and government bond data)
- **E**ach bond priced at pre-selected future times (1 mth, 6 mths, 1yr, 3yrs)
- **R**ecovery rate on default assumed to be 30%
- **V**aR and CVaR computations

The Portfolio of Risky Bonds Demo….

Portfolio Example 3: A Portfolio of Convertibles

A convertible is a bond that can be transformed into an equity holding at some future time(s).

A convertible can have many other ‘bells and whistles’ – can be efficiently evaluated by sophisticated PDE methodology.

Variables: stock price, interest rate (different maturities), volatility, prob. of default (2 predetermined parameters), recovery rate.

The point of this example: the ‘cluster at your finger tips’ allows the fund manager/analyst to aggressively ask future scenario-type questions.

Portfolio Demo 3; Pricing (and sensitivity calculations) a portfolio of convertibles…

Joint demo with ITO 33 (France).

ITO33 The Free Boundary™

Demo Summary

Different portfolio characteristics/requirements lead to different division of work for effective parallel implementation:

Demo 1: Each bond is relatively expensive to evaluate – we chose to have each processor price each instrument independently (minimal communication) with some redundant computation.

Demo2: Instruments are less expensive. We chose to parallelize over the simulations – future time distribution calculations are expensive. Histogram and tail info sent back to master.

Demo 3: Portfolio of convertible bonds. Expense is sensitivity and 'what if' calculations

Conclusion

Many portfolio problems lend themselves to efficient solution on cost-effective clusters comprised of industry standard components. Still, care must be taken on where to parallelize, where to allow redundant calculations,…

It is possible (desirable?) to use the .Net environment and an Excel front end with the core computations done in parallel using web services.

Its clean and effective!

More info, more details?

Email: coleman@tc.cornell.edu