"... Amazon’s real business down the line will be its cloud services. Amazon will be like a book store that sells cocaine out the back door. Books will be just a front to sell storage and cloud computing."

What is Amazon?
- 7 Countries
- More than 84 million active customer accounts
- More than 1.4 million active seller accounts
- Hundreds of thousand of registered associates
- Over 400,000 registered web services developers
- Over 20,000 employees World Wide
- More than 30 Fulfillment Centers worldwide
- 31% of units sold by 3rd party sellers
- 3.9M shipments on peak day in 2007

Amazon’s History in 30 Seconds
- Transformations
  - From single seller to 1M+ sellers
  - From technology consumer to technology provider
  - From single application to platform
  - From web site & database to a massively distributed system
  - From web site to web services
  - From enterprise scale to web scale
Develop
Test
Operate

Undifferentiated heavy lifting
- Hardware costs
- Software costs
- Maintenance
- Load balancing
- Scaling
- Utilization
- Idle machines
- Bandwidth management
- Server hosting
- Storage Management
- High availability

The 70/30 Switch
30% of time, energy and dollars on differentiated value creation

70% of time, energy and dollars on undifferentiated heavy lifting

Livejournal Status

Dream or Nightmare?

- 100% - 1000% increase in popularity?
- Provide 4 nines uptime?
- Survive a complete datacenter failure?
- Survive a router flapping?
- While keeping cost low at the same time?

Web-scale Computing

Scalable Infrastructure that allows applications to meet infinite demand, cheaply and reliably
Scalable
• Increase or decrease capacity in minutes

Cost-Effective
• Low rate, pay-as-you-go

Reliable
• Amazon’s proven infrastructure

Secure
• Multilayer security facilities

Security
Scalability
Availability
Performance
Cost-effectiveness
Incremental Scalability

- Service can be scaled up/down one node at a time.
- Elastic: capable of growing and shrinking on demand.
- Minimal disruption to customer performance.
- Common, fully automated operation.
- Addresses:
  - Different growth paths
  - Fault-tolerance
  - Heterogeneity
  - Operational efficiency

Security
Scalability
Availability
Performance
Cost-effectiveness

Everything fails, All the Time

Failures are Highly Correlated

Systems do not Fail by Stopping
Security  
Scalability  
Availability  
Performance  
Cost-effectiveness

Engineering for Performance at 99.9%

Security  
Scalability  
Availability  
Performance  
Cost-effectiveness

Addressing Uncertainty
• Acquire resources on demand
• Release resources when no longer needed
• Pay for what you use
• Leverage other’s core competencies
• Turn fixed cost into variable

Enterprise Partners
Red Hat, Sun, Oracle, Microsoft
"In a recent test, the Defense Information Systems Agency compared the cost of developing a simple application called the Tech Early Bird on $30,000 worth of in-house servers and software with the costs of developing the same application using the Amazon Elastic Compute Cloud from Amazon.com’s Web Services. Amazon charged 10 cents an hour for the service, and DISA paid a total of $5 to develop an application that matched the performance of the in-house application."