University of Nebraska Omaha

IS&T

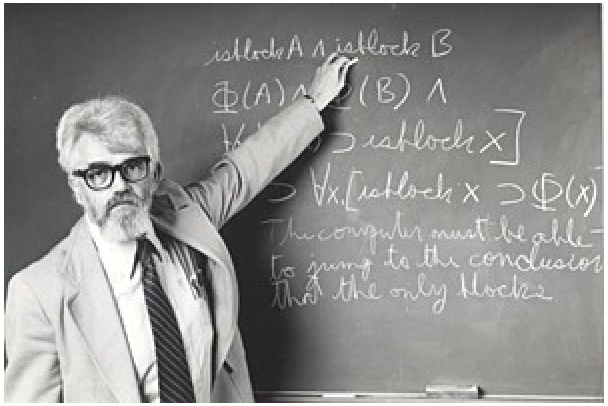
IT Innovation Through Collaboration

Leveraging Distributed Architecture for Cloud Security

Abhishek Parakh

20111013

Working in the Cloud

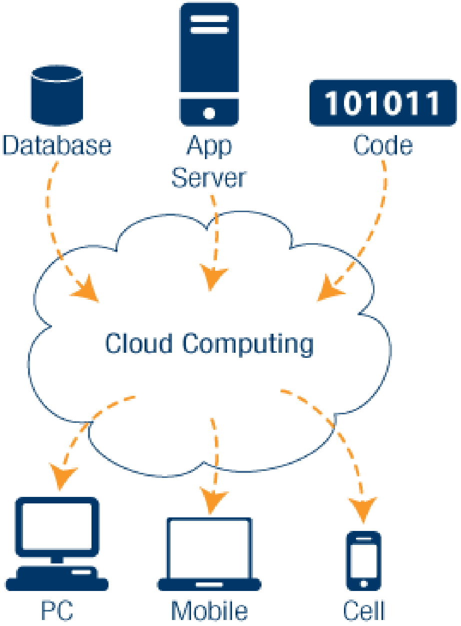


Computation may someday be organized as a public utility.

- John McCarthy, 1960

Working in the Cloud

- Cloud computing is Web-based processing and storage. Software and equipment are offered as a service over the Web.
  - Data and applications can be accessed from any location
  - Data and applications can easily be shared through a common platform
  - Clouds need not be public; companies can introduce private cloud computing solutions

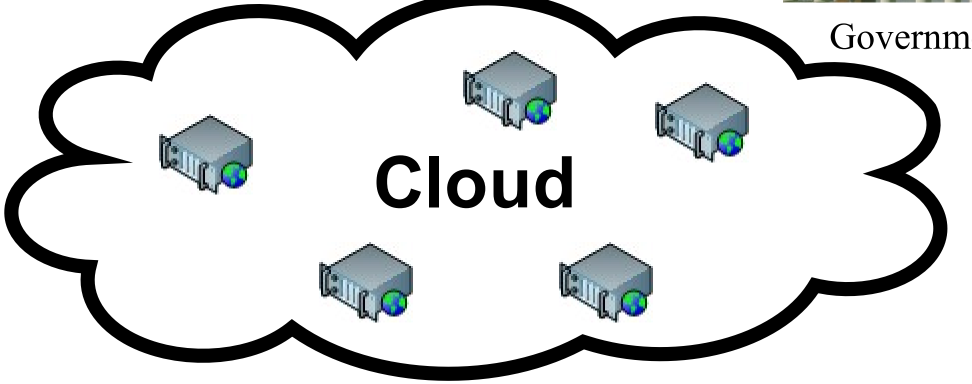


Cost Reduction & Convenience

Small Business

Government Offices

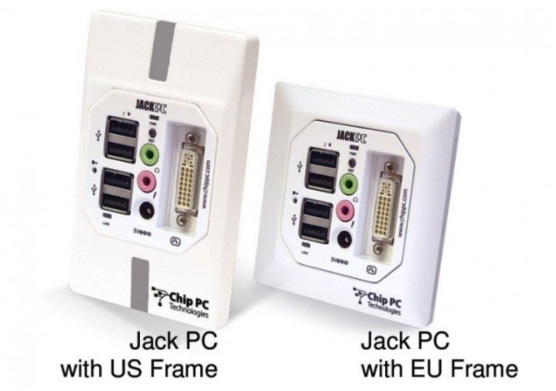
Multinational Corporations

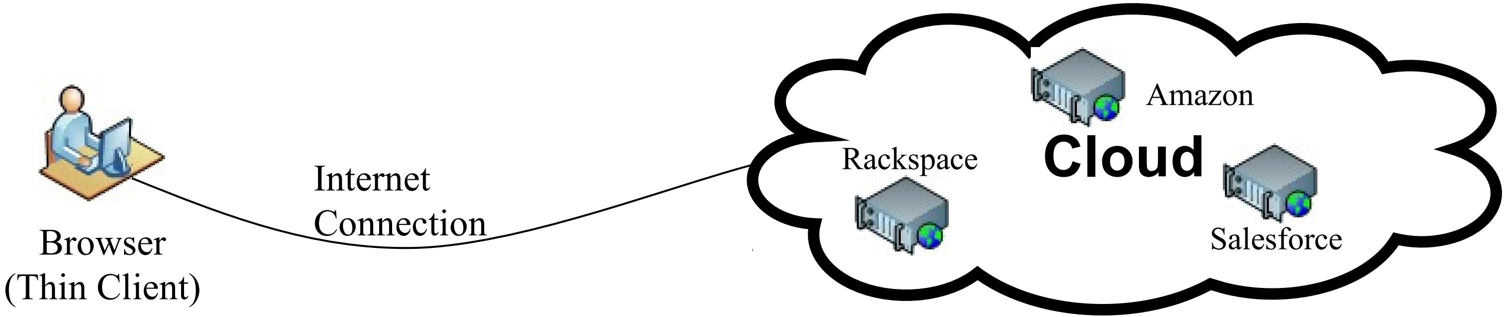


Homes

- Flexible availability of resources
- Opportunity for developers to easily push their applications
- Targeted advertising

The Future of Thin Clients





Working in the cloud

- Cloud Computing for Digital Nomads
  - Cloud computing allows road warriors, telecommuters, and freelancers to connect to their employers using affordable computing tools.
  - Digital Nomads enjoy location independency and the ability to set their own hours, but do not always have job security or benefits.

Three Major Cloud Computing Service Provider Models

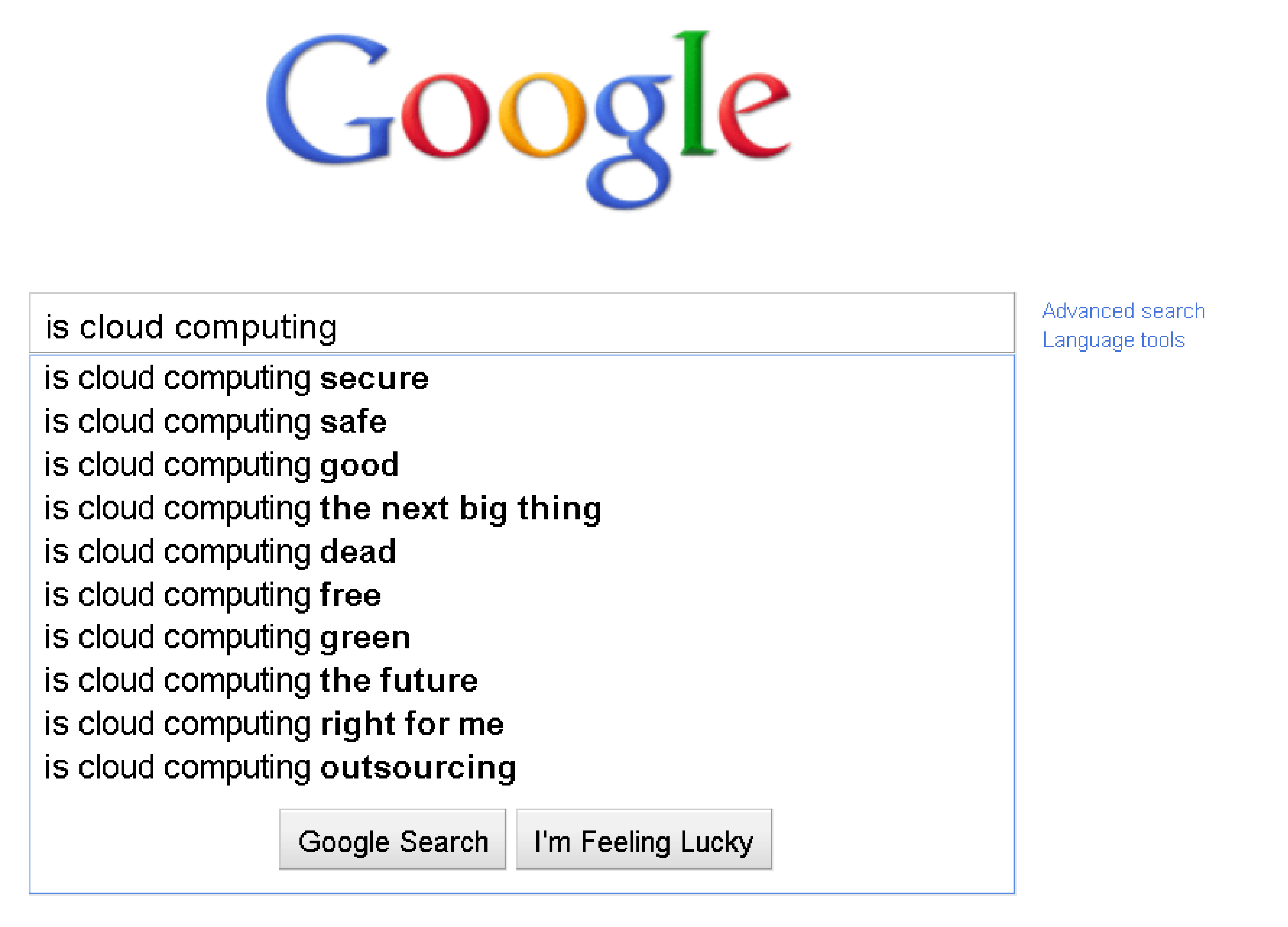
- **Software-as-a-Service** is a model of software deployment in which an application is licensed for use as a service provided to customers on demand. On-demand licensing and use relieves customer of the burden of equipping a device with every application to be used.

- **Platform-as-a-Service:** With the PaaS model, all of the facilities required to support the complete life cycle of building and delivering web applications and services are available to developers, IT managers, and end users entirely from the Internet, without software downloads or installation

- PaaS offerings include facilities for
  - application design, application development, testing, deployment, and hosting,

  as well as application services such as
  - Team collaboration, web service integration and marshaling
  - Database integration, security, scalability
  - Storage, application versioning, application instrumentation
  - And developer community facilitation

- **Infrastructure-as-a-Service** is the delivery of computer infrastructure (typically a virtualized environment) as a service
  - Rather than purchasing servers, software, data center space, or network equipment, clients buy these resources as a fully outsourced service.

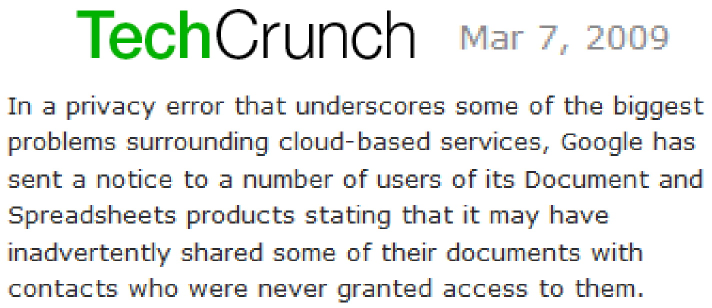


is cloud computing

is cloud computing **secure**
is cloud computing **safe**
is cloud computing **good**
is cloud computing **the next big thing**
is cloud computing **dead**
is cloud computing **free**
is cloud computing **green**
is cloud computing **the future**
is cloud computing **right for me**
is cloud computing **outsourcing**

Cloud Security Incidents

- Cloud-based providers are an attractive target for hackers:
  - In July 2009, a hacker broke in to the personal Web services accounts of Twitter co-founder Evan Williams, his wife, and another Twitter employee, and used that access to steal a number of confidential company documents.
- Multiple levels of Cloud security are needed:
  - Encryption of all communications
  - Server-side security, including regular third-party audits
  - Client-side security, such as firewalls and anti-virus software
  - Client-side password security



TechCrunch Mar 7, 2009

In a privacy error that underscores some of the biggest problems surrounding cloud-based services, Google has sent a notice to a number of users of its Document and Spreadsheets products stating that it may have inadvertently shared some of their documents with contacts who were never granted access to them.

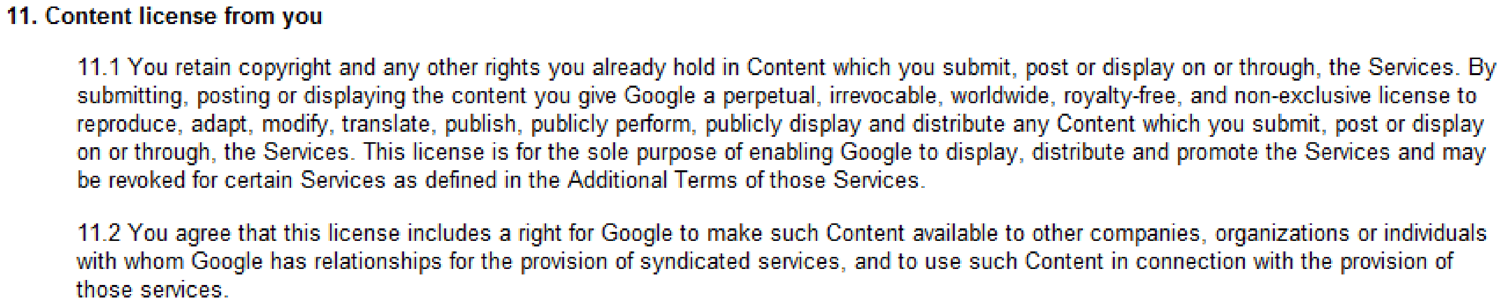
Cloud Security Incidents

- Dropbox’s password nightmare highlights cloud risks
  - Cloud storage site – 25 million users
  - Glitch let visitors use any password to log in customer’s accounts
- Blackberry outages this month
- Neflix accidentally revealed rental histories
- Insurer seeks patient’s Facebook posts
- Sony PlayStation Network hacking

Cloud Security Challenges - Data

- Cloud-based providers need a data storage policy:
  - How is the data supplied to the service housed, protected, shared, manipulated, and disposed of?
  - Who owns the data? Who owns the meta-data?
  - Who has access to the provider’s systems?
  - How can the data be used by parties other than the owner?
  - What is the provider’s data retention policy?

Check out the Google Docs terms of service:



11. Content license from you

11.1 You retain copyright and any other rights you already hold in Content which you submit, post or display on or through, the Services. By submitting, posting or displaying the content you give Google a perpetual, irrevocable, worldwide, royalty-free, and non-exclusive license to reproduce, adapt, modify, translate, publish, publicly perform, publicly display and distribute any Content which you submit, post or display on or through, the Services. This license is for the sole purpose of enabling Google to display, distribute and promote the Services and may be revoked for certain Services as defined in the Additional Terms of those Services.

11.2 You agree that this license includes a right for Google to make such Content available to other companies, organizations or individuals with whom Google has relationships for the provision of syndicated services, and to use such Content in connection with the provision of those services.

Cloud Security Challenges - Availability

- Cloud-based computing requires an availability policy:

What is the provider doing to ensure that client data remains available, even in the event of a natural or human-induced disaster?

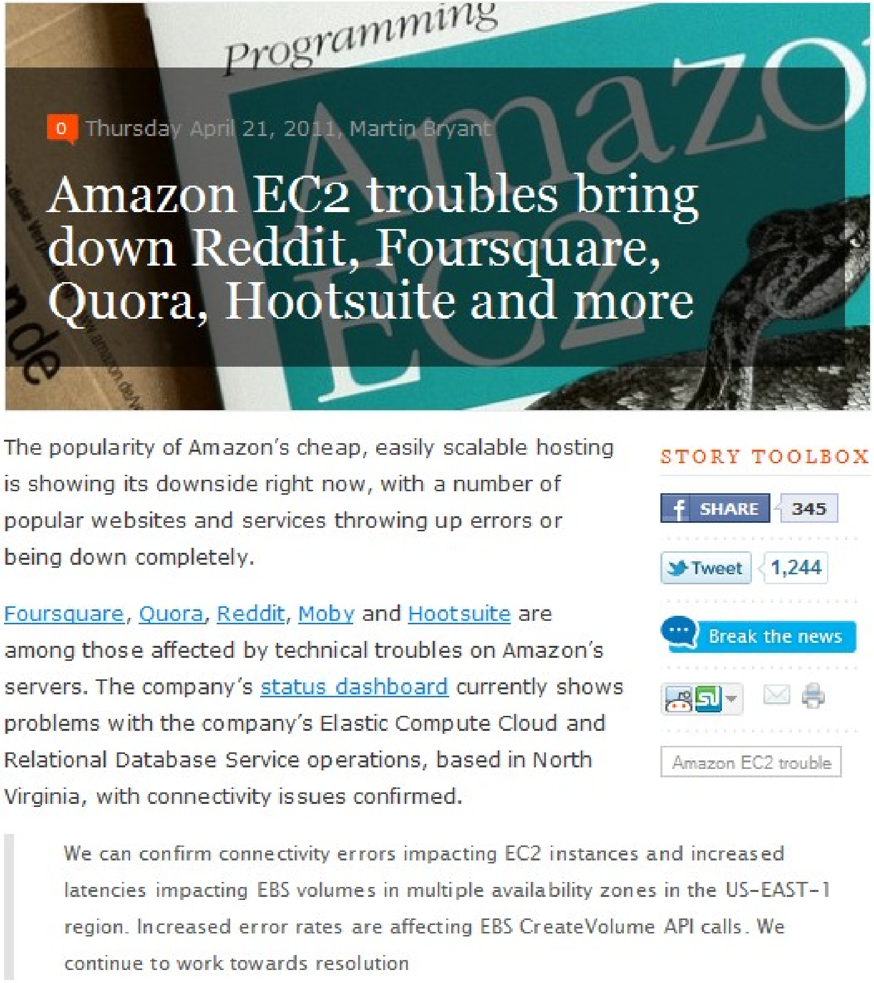
  - Provider back-ups should be frequent
  - Geographic redundancy reduces single-location disasters
  - Are client-side backups possible?
    - Use of open data standards allows data to be extractable or transferrable to other storage locations

Flickr Accidentally Deletes a User's 4,000 Photos and Can't Get Them Back

THE NEW YORK OBSERVER February 1, 2011

Cloud Security Challenges - Availability

TNW
INDUSTRY
04/21/11



0 Thursday April 21, 2011, Martin Bryant

Amazon EC2 troubles bring down Reddit, Foursquare, Quora, Hootsuite and more

The popularity of Amazon's cheap, easily scalable hosting is showing its downside right now, with a number of popular websites and services throwing up errors or being down completely.

Foursquare, Quora, Reddit, Moby and Hootsuite are among those affected by technical troubles on Amazon's servers. The company's status dashboard currently shows problems with the company's Elastic Compute Cloud and Relational Database Service operations, based in North Virginia, with connectivity issues confirmed.

> We can confirm connectivity errors impacting EC2 instances and increased latencies impacting EBS volumes in multiple availability zones in the US-EAST-1 region. Increased error rates are affecting EBS CreateVolume API calls. We continue to work towards resolution

Cloud Security Challenges - Cancellation

- What if your service provider’s business folds or is acquired?
  - Can data and business information be retrieved simply and easily, and in a format useful to the data owner?
  - Will cloud-based data be securely destroyed so others will not be able to access it later?

Cloud Security Challenges - Location

- Location of providers and data centers matter:
  - Local privacy requirements, libel laws, and obscenity regulations may be applied to cloud-based data.
  - If data is stored overseas, does that affect the client's ability to retrieve and produce information in litigation?
  - If data is stored overseas, is it now more accessible to domestic or international government investigators?
  - Does the service agreement limit the jurisdictions in which a client’s data may reside?

Cloud Security Challenges

- Enterprise security is only as good as the least reliable partner, department or vendor.
  - Can you trust your data to your service provider?
- With cloud model – no control over physical security
  - Limited or no knowledge or control of where the shared resources run
  - Exposing your data in an environment shared with other companies could give the government “reasonable cause” to seize your assets because another company has violated the law.

Cloud Security Challenges – contd.

- Storage services provided by one vendor may be incompatible with another vendor’s services should you decide to switch vendors
  - Example: Amazon’s Simple Storage Service (S3) is incompatible with IBM’s Blue Cloud, or Google, or Dell

Cloud Security Challenges – contd.

- Who controls the encryption/decryption keys?
  - Customer / cloud vendor
- Customers want:
  - SSL both ways across the Internet
  - Data encryption when data is at rest
  - Ideally customer must control the encryption/decryption keys

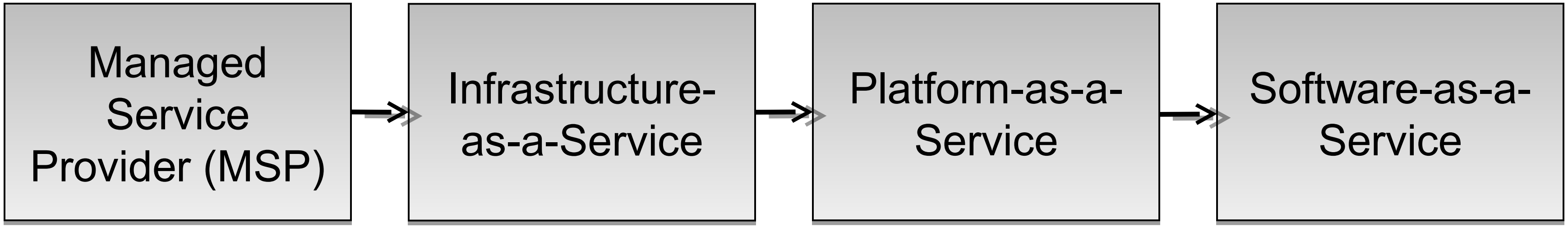
Cloud Security Challenges – contd.

- Data Integrity – assurance that data is identically maintained during any operation (such as transfer, storage, or retrieval)
  - Consistency and correctness
- Data must change only in response to authorized transactions
  - Unfortunately, there are no common standards

Cloud Security Challenges – contd.

- Proper fail-over technology
- Security at data-level so that data is secure wherever it goes
- Compliance standards do not envision compliance in a world of cloud computing
  - Issues of data privacy, segregation and security

Software-as-a-Service Security



Managed Service Provider (MSP) → Infrastructure-as-a-Service → Platform-as-a-Service → Software-as-a-Service

List of security issues which one should discuss with a cloud-computing vendor:

1. Privileged user access
2. Regulator compliance
3. Data location
4. Data segregation
5. Recovery
6. Investigation support
7. Long-term viability

Forensics

- Forensics is used to retrieve and analyze data
  - Responding to an event by gathering and preserving data, analyzing data to reconstruct events, and assessing the state of an event
  - Network forensics include recording and analyzing network events to determine the nature and source of information abuse, security attacks, and other such incidents on your network

- Cloud storage implementations expose a cryptographic checksum or hash (such as the Amazon S3 generation of an MD5 hash) when you store an object

Data Privacy and Governance

- Formal privacy processes and initiatives must be defined, managed, and sustained
- Privacy controls and protection must be an element of the secure architecture design

- Data governance framework should include
  - Data inventory
  - Data classification
  - Data analysis
  - Data protection
  - Data privacy
  - Data retention/recovery/discovery
  - Data destruction

Security as a Service

- Managed Security Service Providers (MSSP)
  - Dominated the outsourced hosting from mid-90s to early 2000’s
- MSSP is an ISP (Internet Service Provider) that provides
  - Network security management
  - Security information management
  - Security event management
  - Virus blocking, spam blocking, intrusion detection, firewalls and VPN (virtual private network) management
  - Handle system changes, modifications and upgrades
- MSSP -> Security as a Service

Network Threats

- DNS Attacks
- Sniffer Attacks
- Distributed Denial of Service Attack
- IP Spoofing
- Malware: Viruses and Worms

Data Storage Security

- Security in data transit
  - Public key cryptosystems

However,

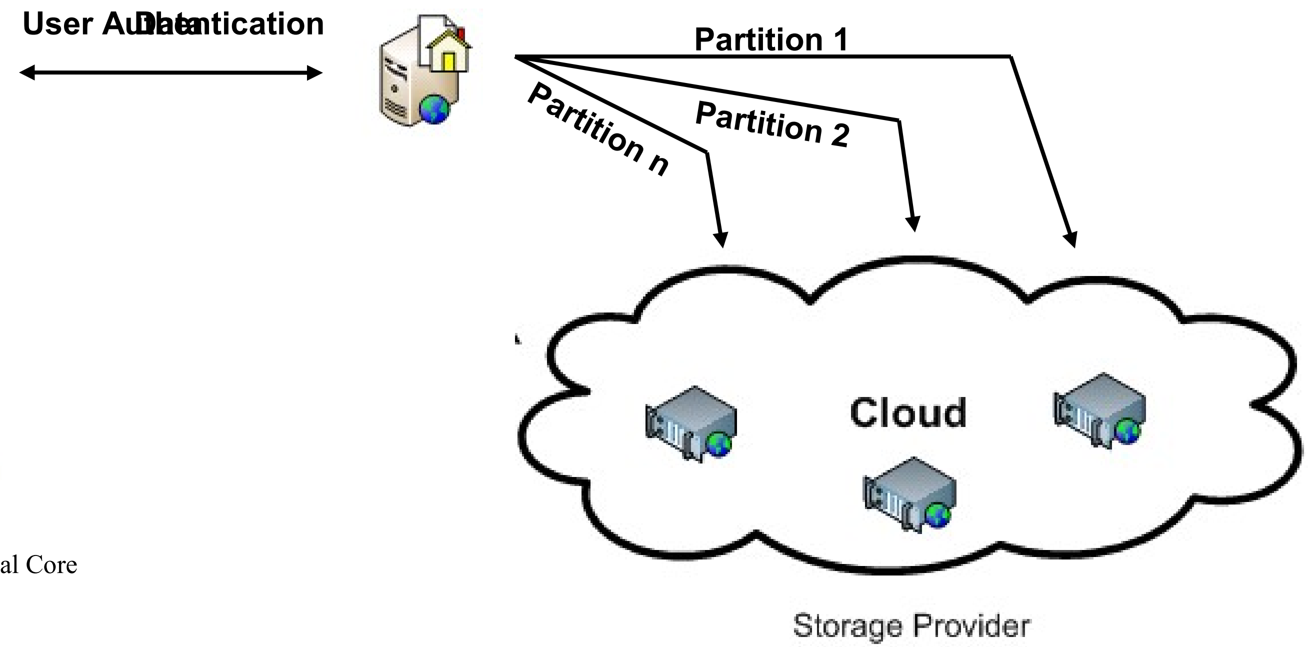
- Large number of data breaches happen on stored data, due to,
  - Internal breaches
  - Unscrupulous employees
  - Weak server passwords
  - Un-patched operating systems

Implicit Security

- Move away from (traditional) key based encryption systems
- Use cryptographic data partitioning
- Adopt a distributed architecture
- Make multiple parties responsible

We Are Going To Be Looking At

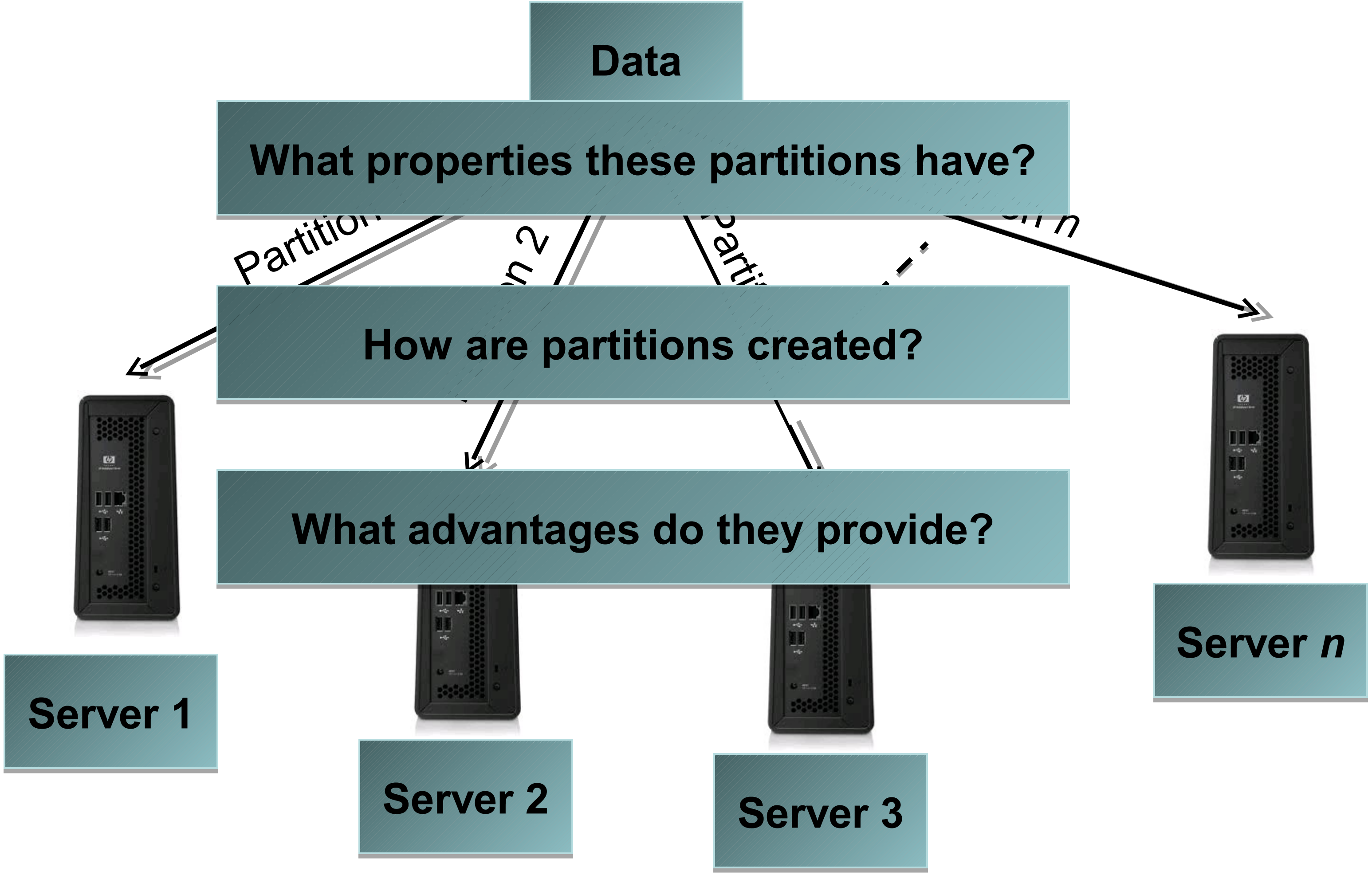
- Secure Data Storage
  - Assume that we are using Online Storage Providers for data storage



Motorola Atrix – 2 GHz Dual Core

- Distributed Architecture
- Data Partitioning

Distributed Solution Data Partitioning



Data

What properties these partitions have?

How are partitions created?

What advantages do they provide?

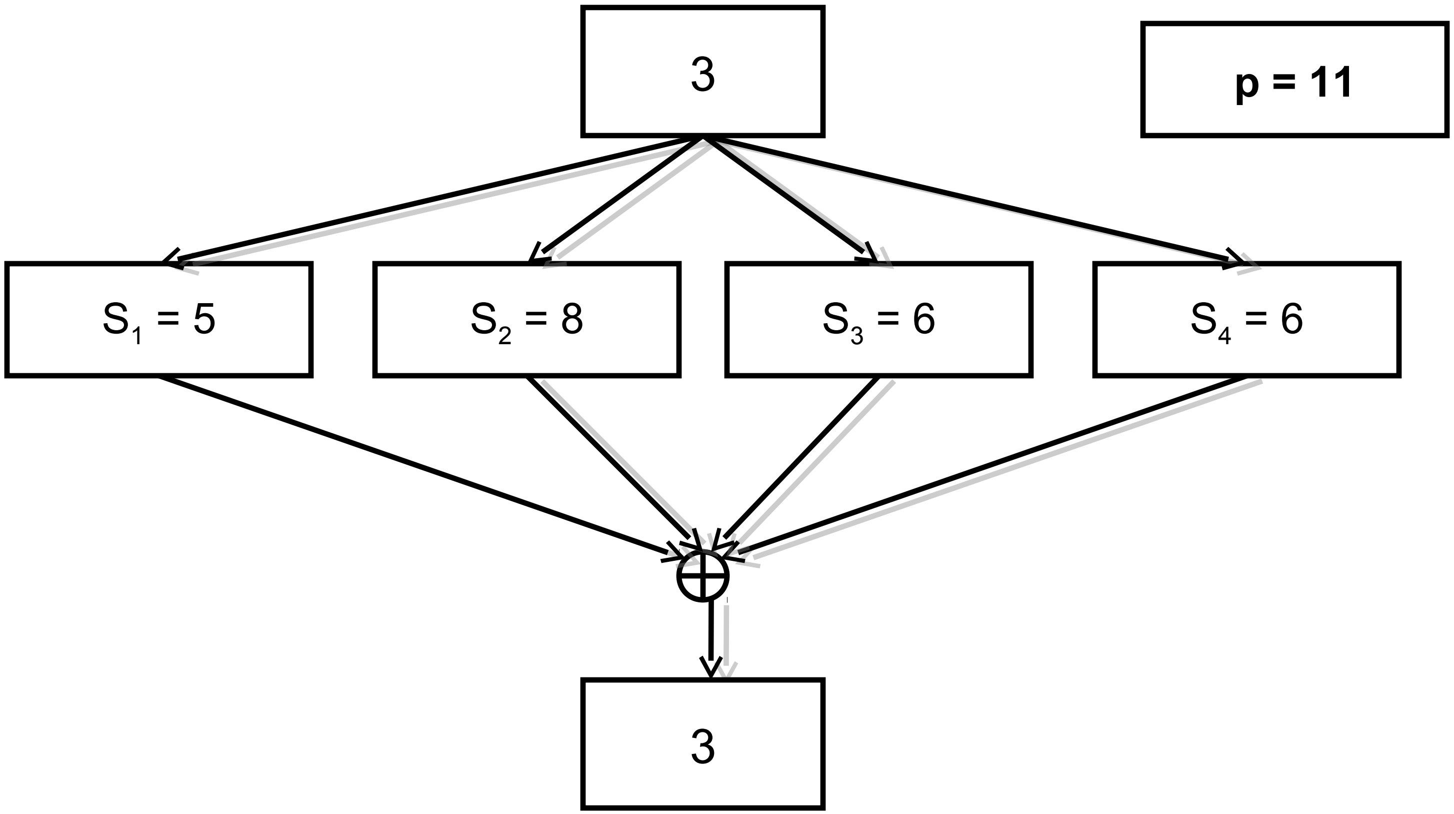
Server 1

Server 2

Server 3

Server *n*

Data Partitioning



3

p = 11

S_1 = 5

S_2 = 8

S_3 = 6

S_4 = 6

$\oplus$

3

- Property:
  - No single partition reveals any information about the data until all of them are added together
  - Drawback: All partitions are needed to recreate the data: No redundancy

Security Level – Highest

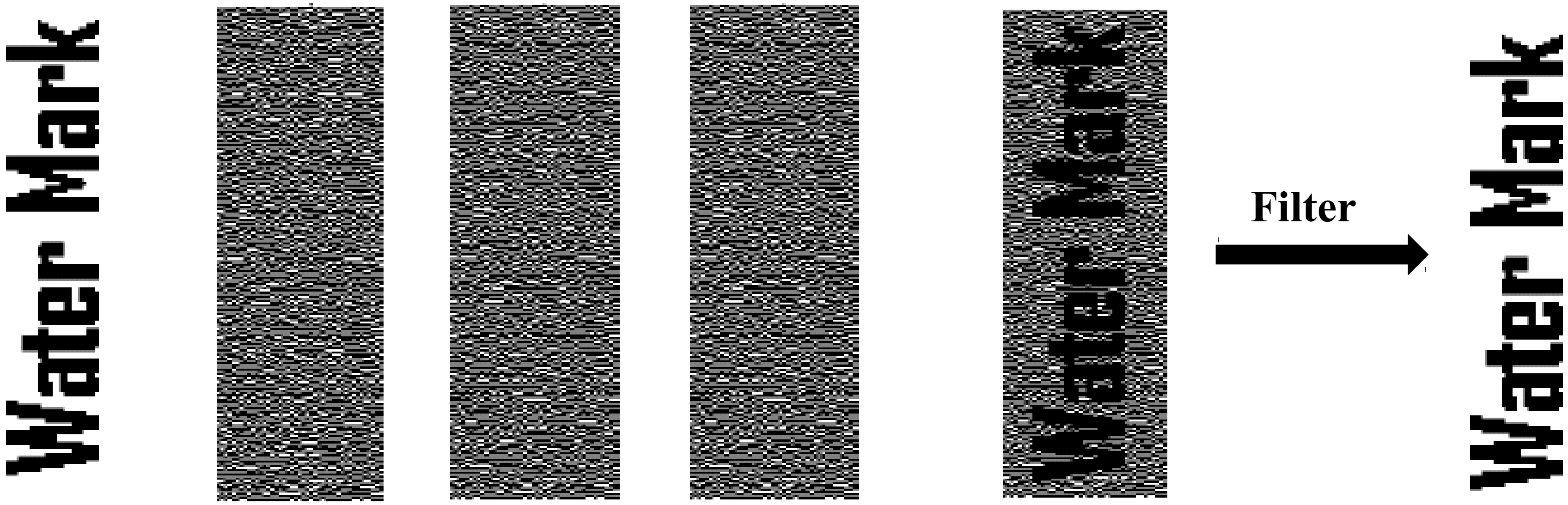
Parameters	
Number of data items encoded	1
Total number of partitions	n
Required number of partitions	k
Redundancy factor	n-k
Storage space per data item	n x data size

Obtaining	**Amount of information revealed**
1 partition	none
2, 3,…, k-1 partitions	none
k partitions	all

Cost v/s Security

- In general we call these schemes (k,n) data partitioning schemes, n ≥ k
  - k = minimum number of partitions required for data reconstruction
  - n-k = redundant pieces
- Large k = better security
- Large n-k = better reliability
- But
  - Large n-k = more points of vulnerability
  - Large n = higher communication and storage cost

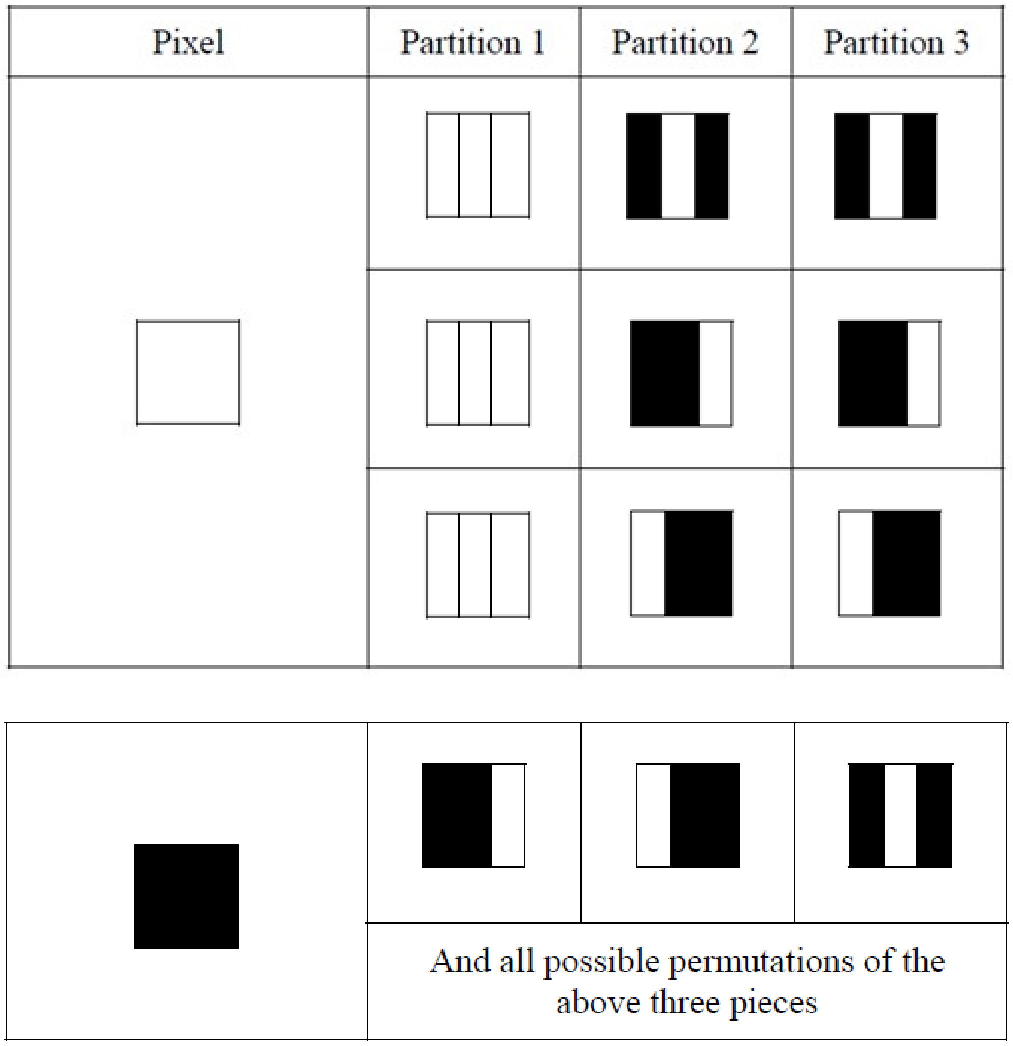
Image Partitioning



Three random looking partitions for the original Water Mark image

Result of overlapping any two partitions

Partitioning Images: one pixel at a time



Pixel

Partition 1

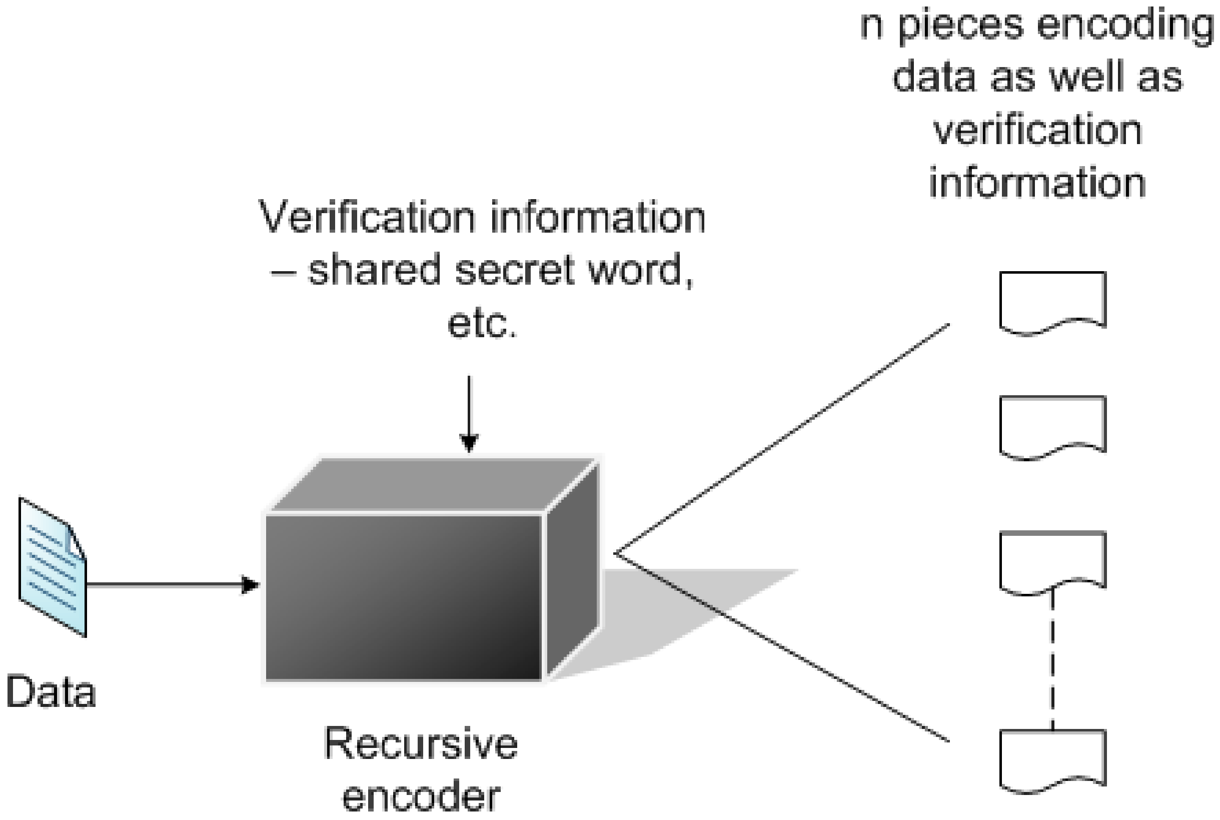
Partition 2

Partition 3

And all possible permutations of the above three pieces

Applications

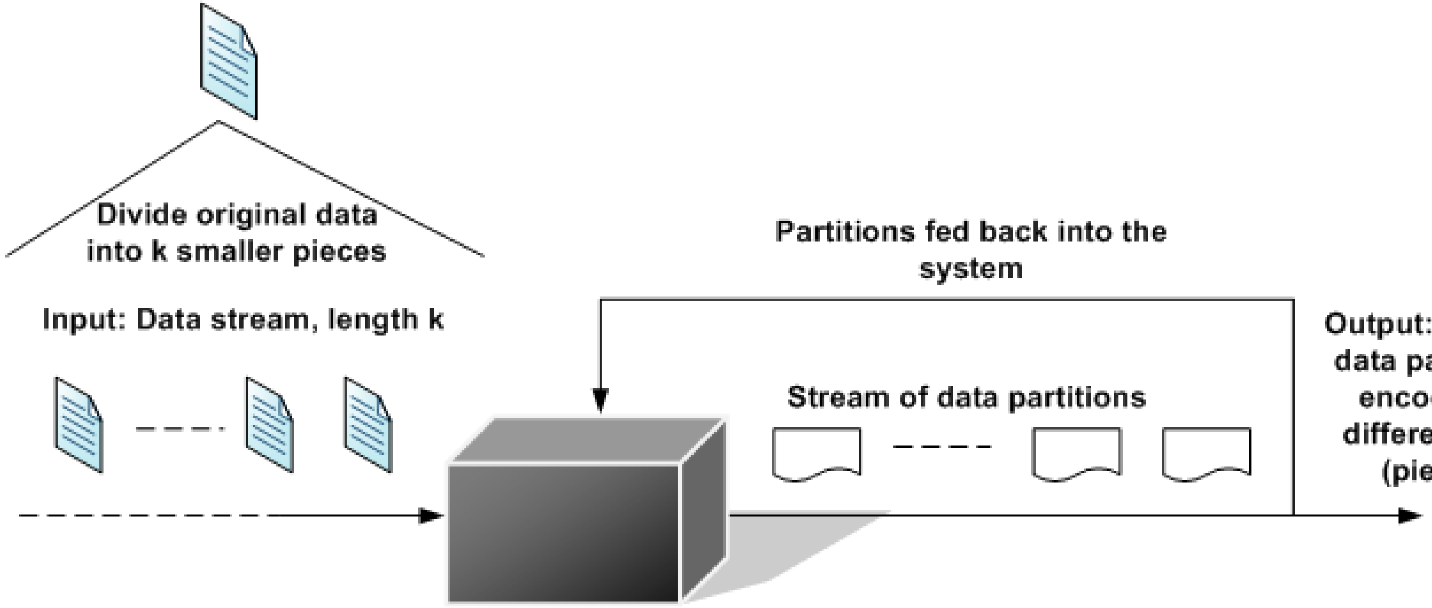
- Recursion
  - Hidden verification information – for self verification of regenerated data (or data)



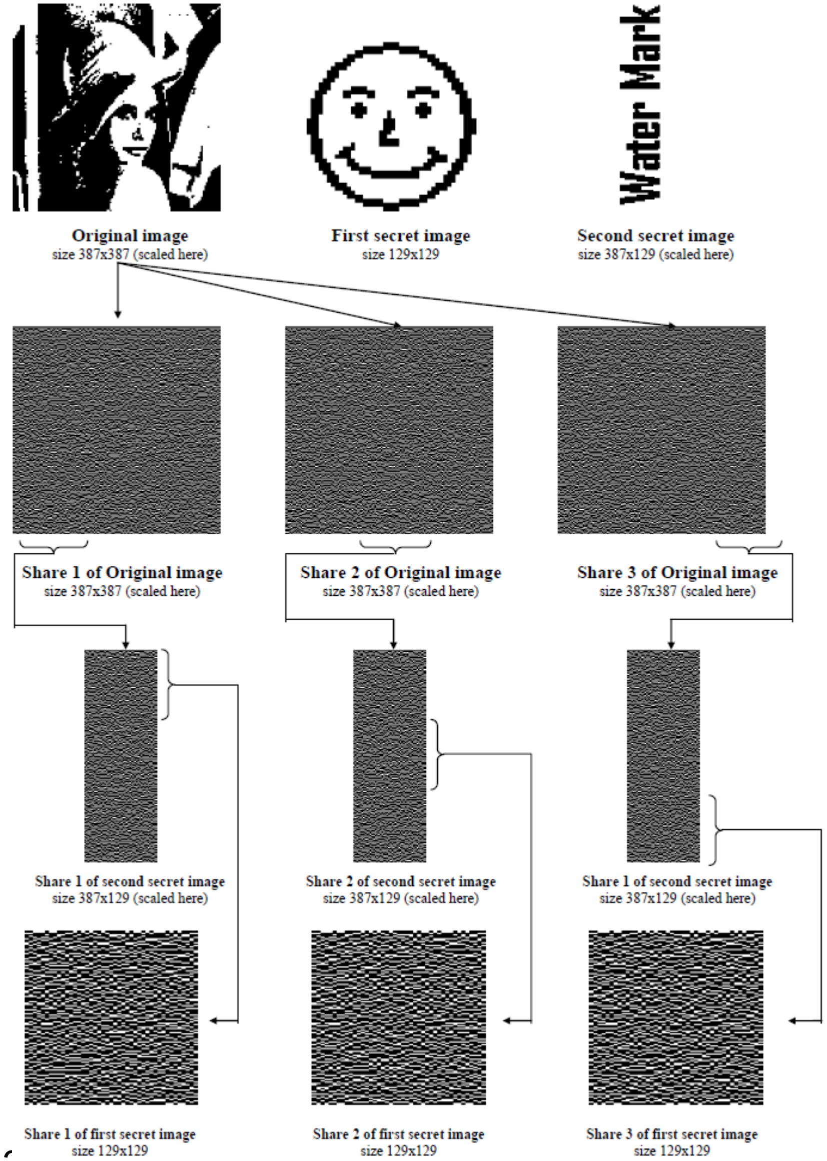
n pieces encoding data as well as verification information
Verification information – shared secret word, etc.
Data
Recursive encoder

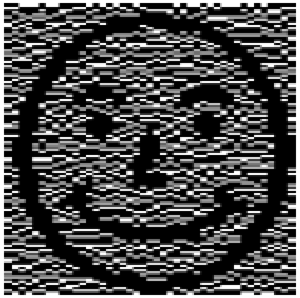
Applications

- Recursion
  - Information dispersal or multi-secret sharing


Divide original data into k smaller pieces
Input: Data stream, length k
Partitions fed back into the system
Stream of data partitions
Output: Set of n data partitions encoding k different data (pieces)

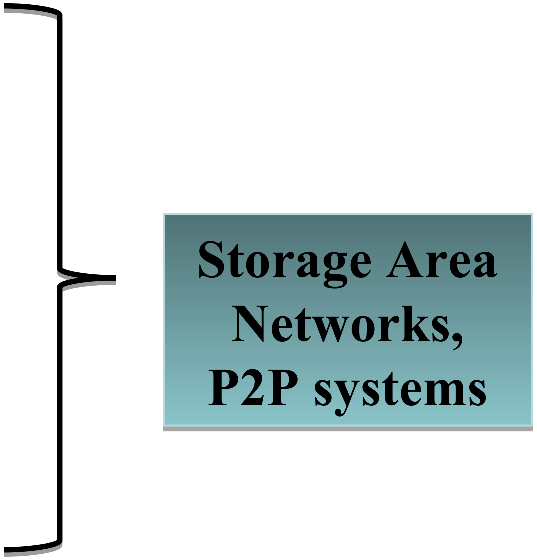
Recursive Hiding in Image Partitioning


Original image
size 387x387 (scaled here)
First secret image
size 129x129
Water Mark
Second secret image
size 387x129 (scaled here)
Share 1 of Original image
size 387x387 (scaled here)
Share 2 of Original image
size 387x387 (scaled here)
Share 3 of Original image
size 387x387 (scaled here)
Share 1 of second secret image
size 387x129 (scaled here)
Share 2 of second secret image
size 387x129 (scaled here)
Share 1 of second secret image
size 387x129 (scaled here)
Share 1 of first secret image
size 129x129
Share 2 of first secret image
size 129x129
Share 3 of first secret image
size 129x129



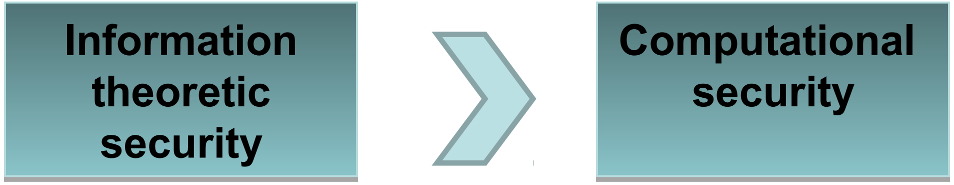
Data Storage Systems Available That use Data Partitioning

- Delta-4
- Intermemory
- Oceanstore
- Farsite
- E-Vault
- PASIS
- Publius



Storage Area Networks, P2P systems

➢ Before using any of the above providers user must determine the level of security they provide



Information theoretic security

Computational security

Holy Grail for Cloud Security

- Homomorphic Encryption
  - An encryption system that allows computations to be performed on encrypted data
    - By enabling both multiplication and addition of encrypted data
  - Subsequent decryption yields the expected result in plaintext

Questions?

Thank You!

Contact Information

Abhishek Parakh

Assistant Professor

School of Interdisciplinary Informatics

College of Information Science and Technology

University of Nebraska at Omaha

aparakh@unomaha.edu

402-554-3161

PKI 177D