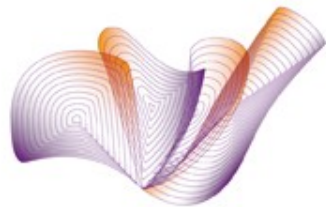


Cloud- and Peer-to-Peer Storage

End-user considerations and product overview

4/3/2010

Arjan Peddemors



novay

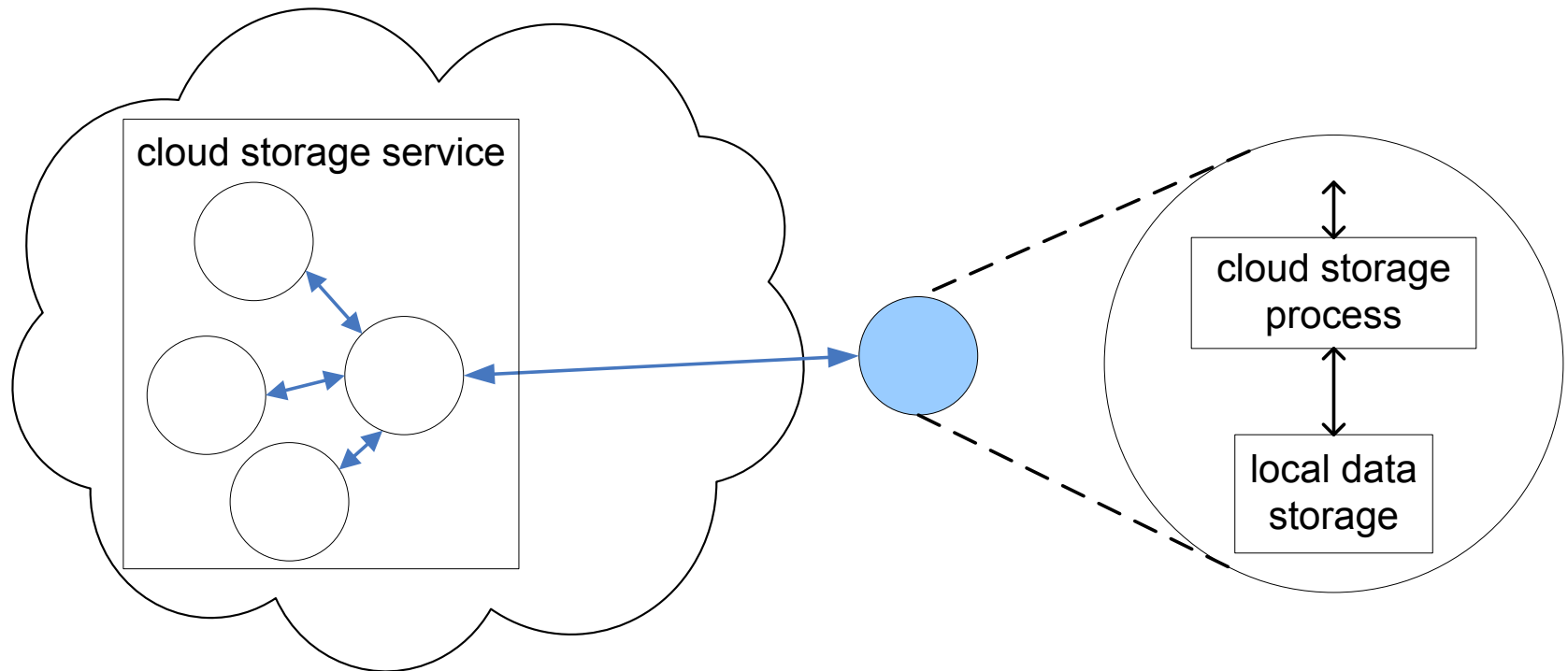
SURF

NET

Objectives

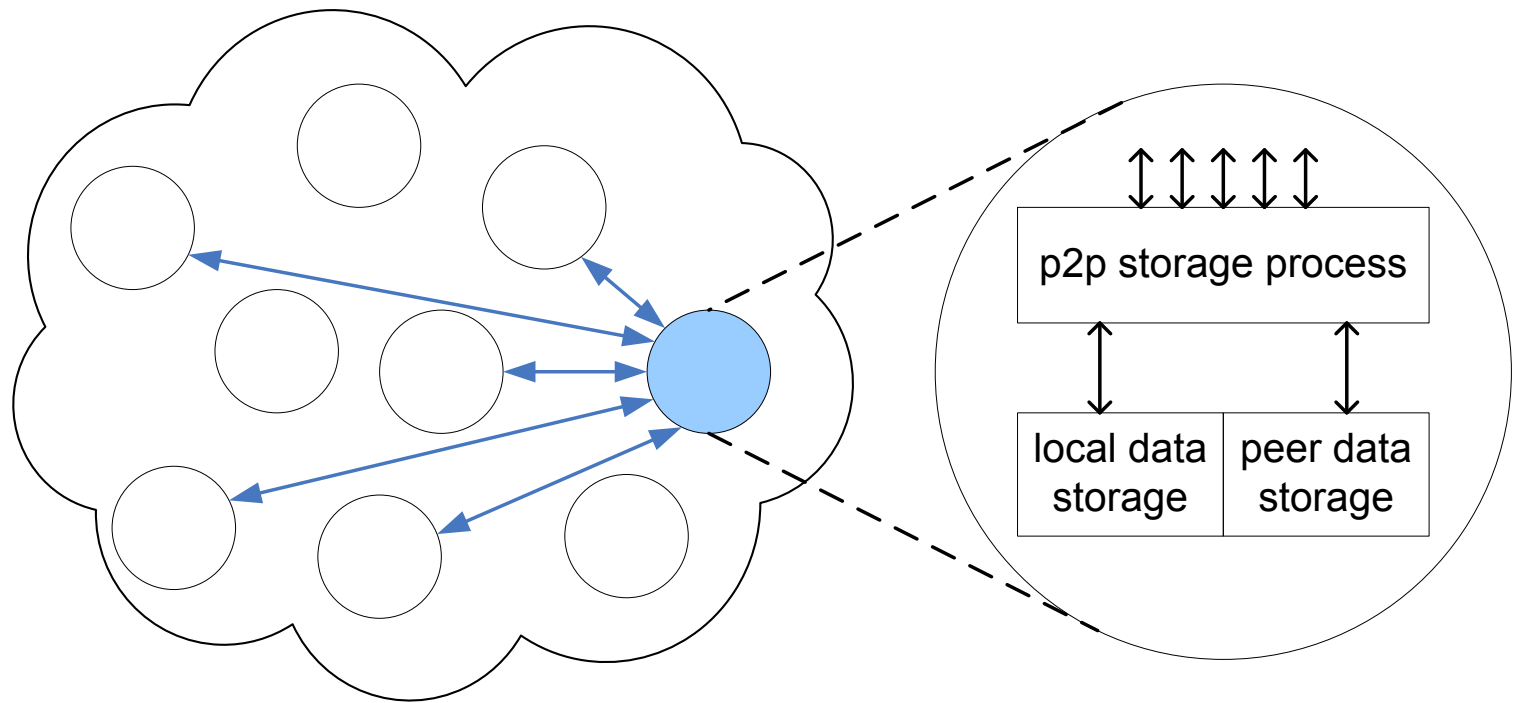
- Get overview of existing cloud storage and P2P storage *concepts and products**
 - outline of basic principles underlying cloud- and P2P storage technologies
 - product scan
 - scenario description
- Focus on end-users (researchers, staff, students)

Overview Cloud Storage



Overview

Peer-to-Peer (P2P) Storage



Overview

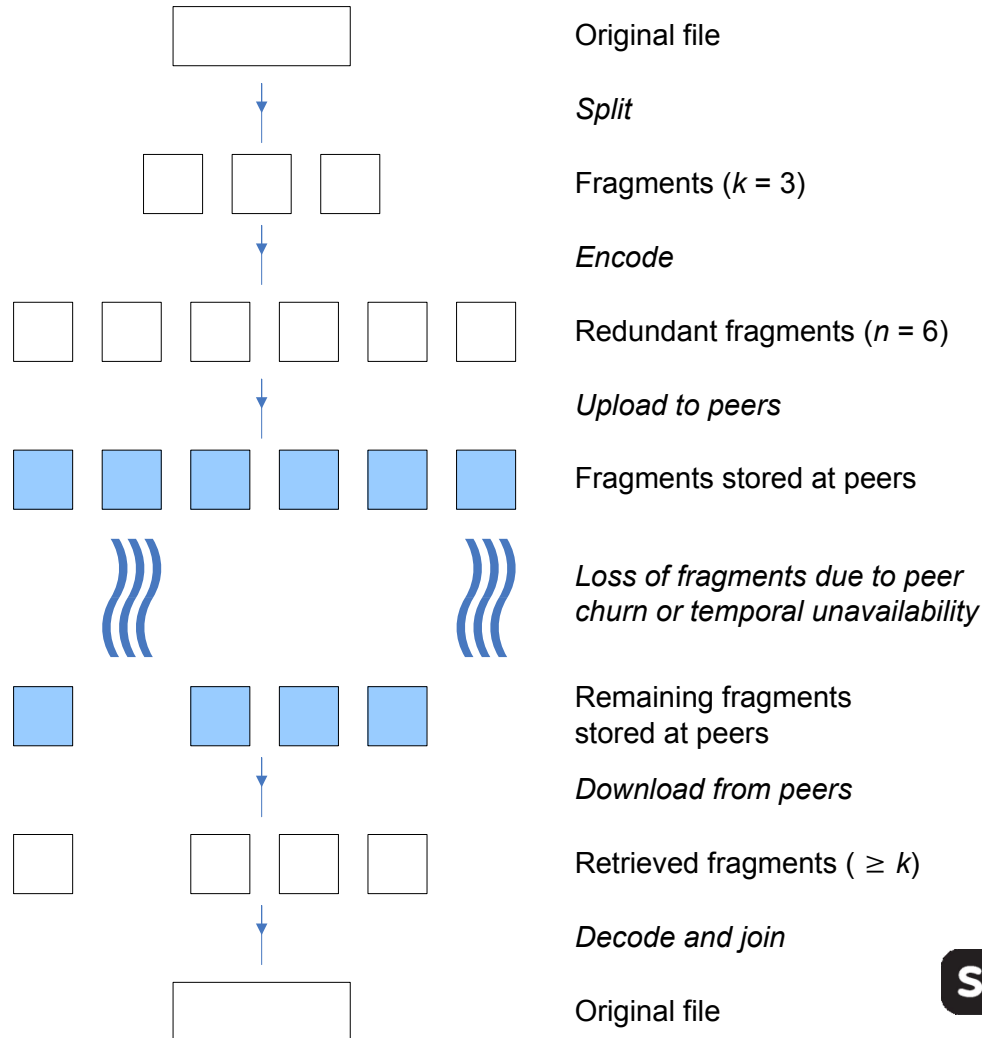
Application Domain

- Cloud storage / P2P storage domain is currently highly dynamic
 - many new products
 - quickly increasing number of (end-)users
 - new technologies being applied
- Future developments driven by
 - fast internet connections for (end-)users
 - cheap storage hardware
 - uptake of cloud computing in general

Basic principles

- Redundancy
- Erasure coding
- Security
- Asymmetric read / write
- Peer selection
- Fairness and quota
- Versioning
- Efficient transmission

Erasure Coding



Original file

Split

Fragments ($k = 3$)

Encode

Redundant fragments ($n = 6$)

Upload to peers

Fragments stored at peers

Loss of fragments due to peer churn or temporal unavailability

Remaining fragments stored at peers

Download from peers

Retrieved fragments ($\geq k$)

Decode and join

Original file

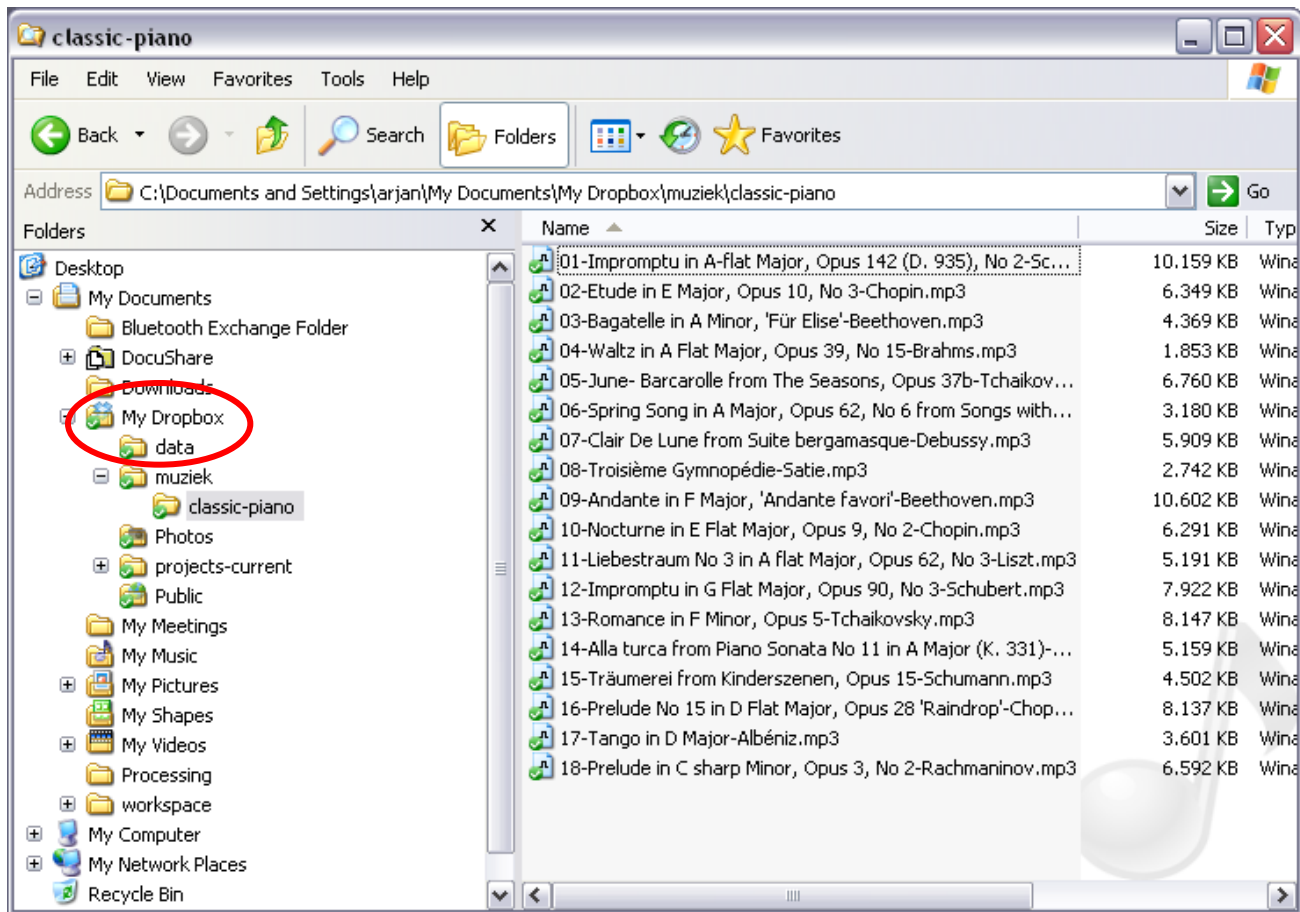
Erasure Coding

- Compared to replication, erasure codes deliver a higher probability of successful data restore
- Example calculation
 - peer network consisting of 10000 peers
 - 10% of peers is offline on average
 - replication factor of 2 (file consumes twice the original size at peers)
 - 2 replicas: probability of availability of at least one replica is 0.99 (2 nines)
 - 16 fragments at peers: probability of availability of at least 8 fragments is 0.99994 (four nines)

Product scan

- Research projects / products
- Simple Tools
- Cloud Storage products
 - Amazon S3
 - Dropbox
 - Microsoft Live Mesh / Live Skydrive
 - Adrive
 - Mozy
- P2P Storage products
 - Wuala
 - Tahoe
 - CrashPlan
 - Cucku
 - PowerFolder

Amazon S3 / Dropbox (cloud)

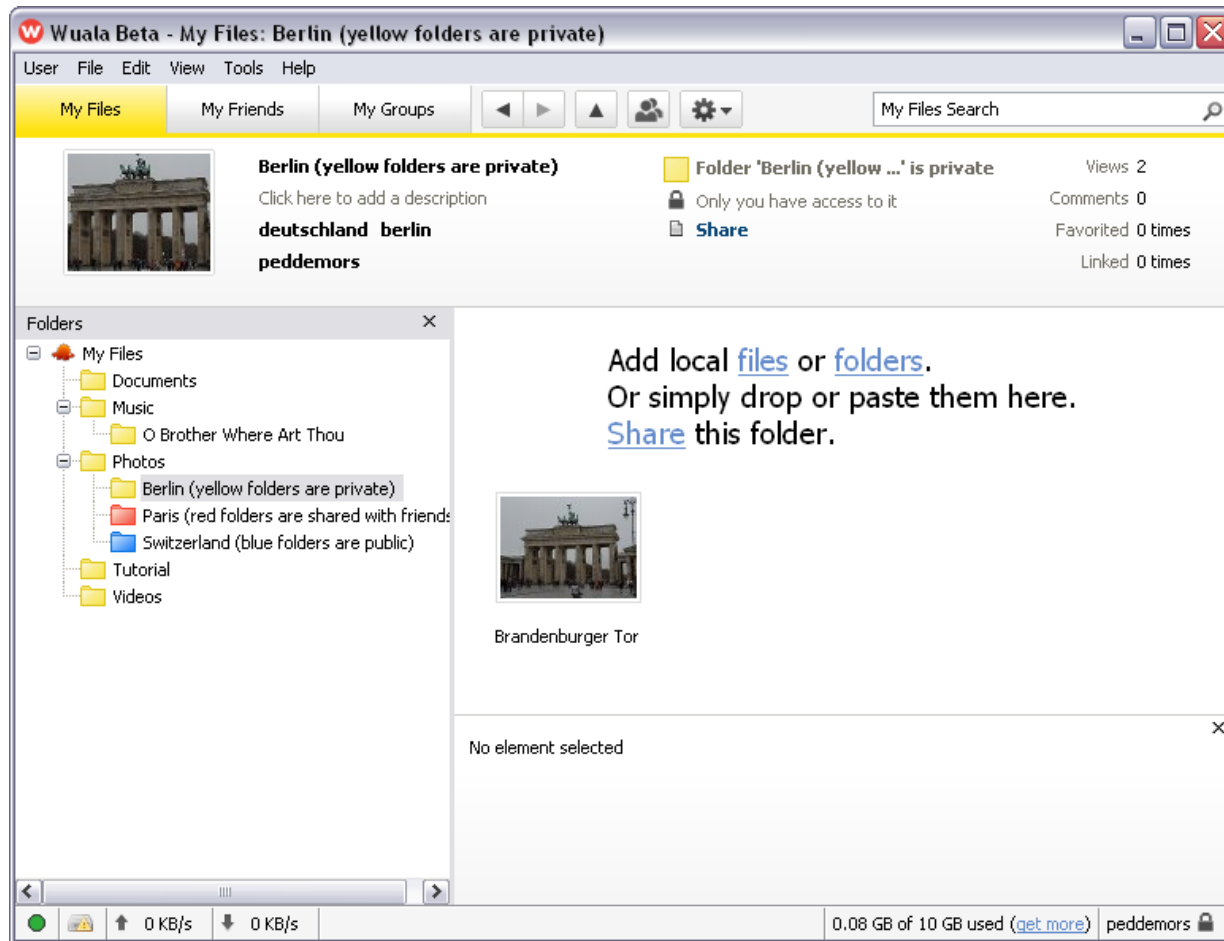


Amazon S3 / Dropbox (cloud)

The screenshot shows the Dropbox web interface in a browser window. The address bar displays the URL <https://www.dropbox.com/home#/muziek/classic-piano/>. The page features the Dropbox logo and navigation tabs for 'Getting Started', 'Files', 'Recent Events', 'Sharing', and 'Help'. The main content area shows the path 'My Dropbox >> muziek >> classic-piano' and a search bar. Below this, there are buttons for 'Upload', 'Create folder', 'Share this folder', and 'More actions'. A table lists files with columns for 'Name', 'Size', and 'Modified'. The file '02-Etude in E Major, ...-Chopin.mp3' is selected, and a context menu is open over it, showing options like 'Previous versions', 'Download file', 'Move...', 'Rename...', 'Copy file to...', and 'Delete...'.

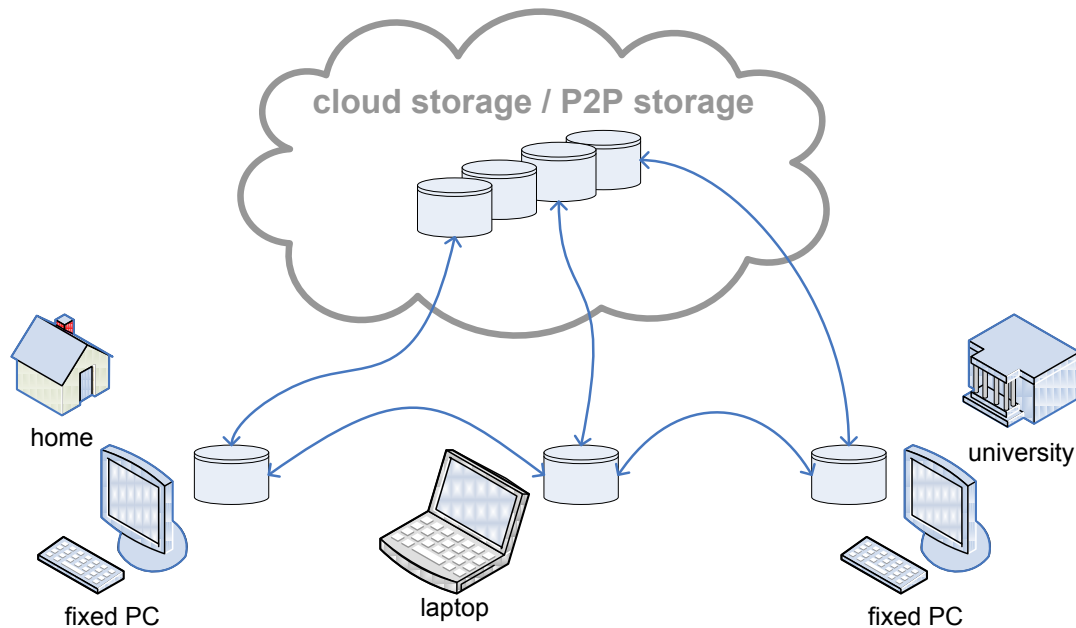
Name	Size	Modified
Parent folder		
01-Impromptu in A-Fla...chubert.mp3	9.92MB	Yesterday 4:04PM
02-Etude in E Major, ...-Chopin.mp3	6.2MB	Yesterday 4:04PM
03-Bagatelle in A Min...ethoven.mp3	4.27MB	
04-Waltz in A Flat Ma...Brahms.mp3	1.81MB	
05-June- Barcarolle f...ikovsky.mp3	6.6MB	
06-Spring Song in A M...elssohn.mp3	3.11MB	
07-Clair De Lune from...Debussy.mp3	5.77MB	
08-Troisième Gymnopédie-Satie.mp3	2.68MB	
09-Andante in E Major...ethoven.mp3	10.35MB	

Wuala (P2P)



Scenario

- Doctoral student working on research data / thesis at multiple locations



Scenario

- Comparing
 - Simple tool (Unison)
 - Cloud storage service (Dropbox)
 - P2P storage service (Wuala)
- Aspects
 - Preparation / installation
 - Regular activities and maintenance
 - Recovery

Conclusions

- Cloud / P2P storage products interesting for SURFnet end users
- Domain is new and developing quickly
- Many product variants (targeting different usage)
- Subscription versus bringing in own resources
- Two types of P2P storage products
 - ‘full’ P2P system scaling to many peers
 - ‘social’ P2P systems
- Cloud storage services more mature than P2P
- Many products have both Cloud and P2P features