

**KOST** Koordinationsstelle für die dauerhafte Archivierung  
elektronischer Unterlagen

---

Ein Gemeinschaftsunternehmen von Schweizer Archiven

# Pilotprojekt *'Speicherbox'*

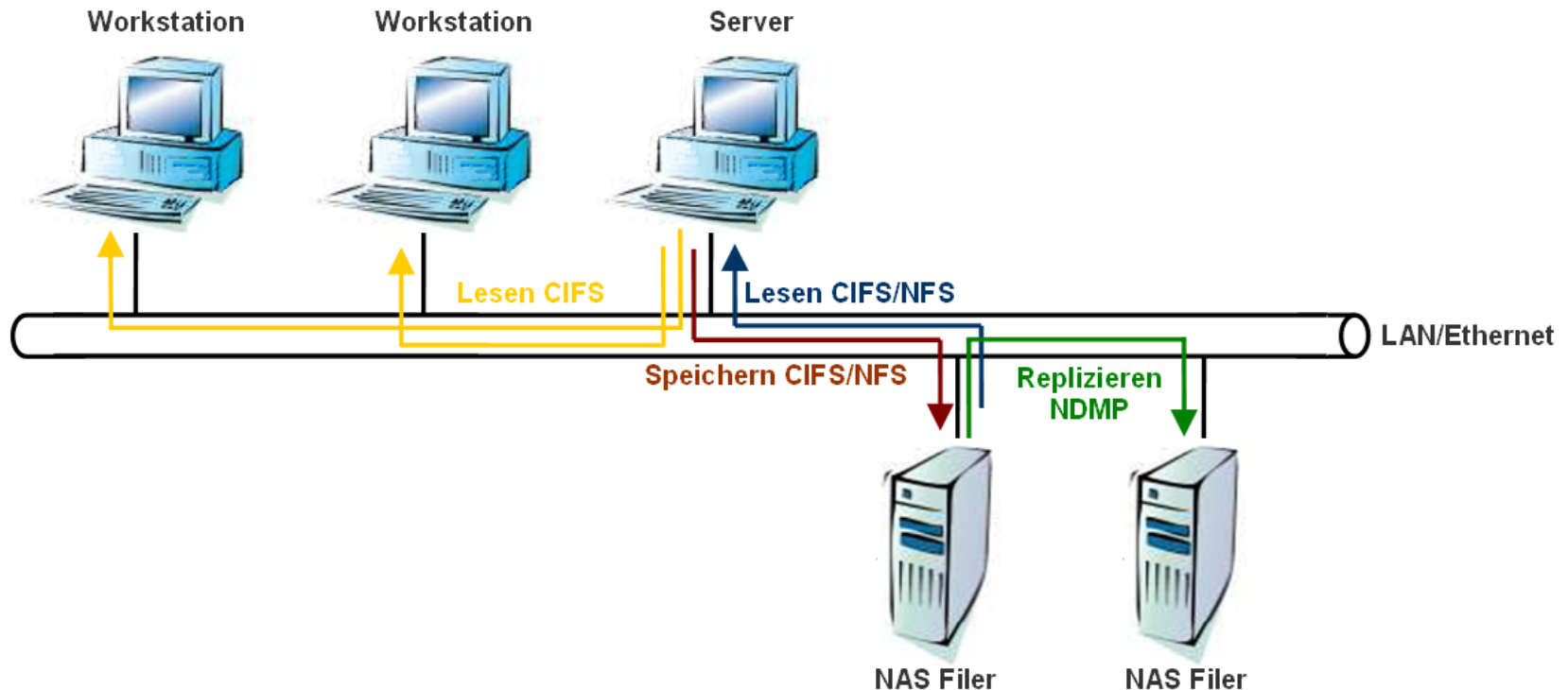
## Technische Umsetzung

Abschlusspräsentation  
Bern, 25.06.2009

# Projektplanung

1. Detailspezifikation der Lösung
2. Budgetplanung
3. Evaluation von Hard- & Software
4. Testen der gewählten Hardware
5. Entwicklung der Softwarelösung
6. Implementierung im Staatsarchiv Bern
7. Testen und Abnahme

# Architektur der Lösung



*Lokale Speicherbox im Archiv mit automatischem Replizieren der Daten*

# Detailspezifikation

- Die Speicherboxen sollen als *Kompaktsystem* realisiert sein.
- Der Zugriff auf die gespeicherten Daten soll via Netzwerk stattfinden (*NAS Network-Attached Storage*).
- Als dateibasierender Dienst muss *SMB/CIFS* angeboten werden.
- Datensicherheit muss mit mind. *RAID Level 5* garantiert werden.
- Zur *Fernüberwachung* kann SNMP oder E-Mail eingesetzt werden.
- Eine *Redundanz* im Bereich RAID/Festplatten-Kontroller, Stromversorgung und Netzwerkkarte ist nicht notwendig.
- *Server-Speicherplatten* für den 7 x 24h Betrieb sind zwingend.
- Eine *USV* sorgt dafür, dass die Speicherbox bei einem Stromausfall sicher heruntergefahren wird.
- Eine *Storage Management Software* soll alle Daten von der einen Speicherbox automatisch auf die zweite Speicherbox *replizieren*.
- Die *Software* soll periodisch die *Integrität* der Daten überprüfen.

# Budget & Evaluation

## *Budget:*

Anzahl	Produkt	Stückpreis	Total
2	Thecus N5200 PRO	1'239.90	2'479.80
12	Seagate Barracuda ES.2 1TB/32MB SATA	272.00	3'264.00
2	APC Smart-UPS 1000VA USB & Serial 230V	549.00	1'098.00
0	EMC Insignia RepliStor SMB Edition 6.1	1'253.70	0.00
1	Verbrauchsmaterial	200.00	200.00
	Mehrwertsteuer	7.6%	535.10
Gesamttotal Hardware/Software (plus / minus 20%)			7'576.90

## *Evaluation:*

Die Evaluation ergibt, dass *EMC RepliStor* nur zusammen mit einem dedizierten Fileserver betrieben werden kann.

Der Entscheid fällt zugunsten von *robocopy* zusammen mit einer Scriptlösung.

# Testbericht zur Hardware

- Der Testbericht zeigt, dass die gewählte NAS Box *RAID 5* korrekt implementiert (RAID 6 oder 10 sind nicht zu empfehlen).
- Der *Stromverbrauch* hält sich in Grenzen und beträgt 84 Watt, 6 Watt gehen auf das Konto der USV.
- Der Datendurchsatz ist hinreichend und die Anbindung an *Microsoft Active Directory*<sup>®</sup> erfolgt problemlos.
- Die Webapplikation zur *Administration* funktioniert perfekt.
- Die *Seagate Barracuda ES.2 1*-Platten melden nach kurzer Laufzeit *reallocated sectors*. Das Problem scheint bekannt zu sein und mit der Grösse der Platten zu tun zu haben (1TB). Weitere Abklärungen sind im Gang.

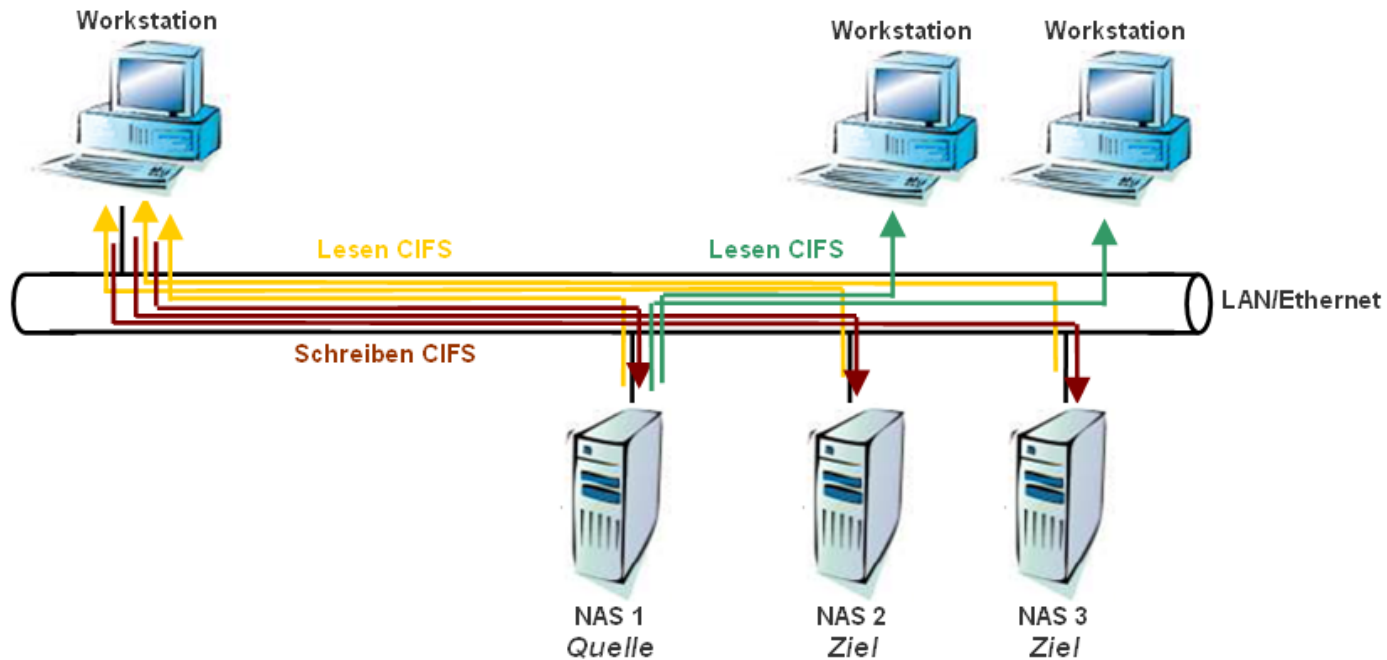
# Softwarelösung

Die Anforderungen aus der Spezifikation werden durch eine Scriptlösung auf der Grundlage von *robocopy.exe* von Microsoft realisiert:

- Die Scriptlösung *autoverify repliziert* mit *robocopy.exe* alle Daten von der einen Speicherbox sicher und automatisch auf weitere Speicherboxen.
- Die Lösung prüft periodisch die *Integrität* der Daten durch Vergleich der Daten auf den Speicherboxen.
- Replikation und Integritätscheck können *fortlaufend*, aber auch periodisch erfolgen.
- Die *System- und Netzwerklast* kann individuell festgelegt werden und weniger als 5% betragen.
- Die Scriptlösung kann auch auf einer *Workstation* installiert werden, die regelmässig abgeschaltet wird.

# Softwarelösung

Topologie: eine Arbeitsstation mit *autoverify*, drei NAS und zwei Arbeitsstationen mit lesendem Zugriff auf ein NAS:





# Softwarelösung

```
c:\cmd
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

N:\KOST\Projekte\Sauer_les_CD\Speicherbox_StABE\autoverify>autoverify.exe

Usage :: autoverify.exe source destination;dest2;dest3 [options]
Source  :: Source Directory (drive:\path or \server\share\path)
Destination :: Directory list to compare and copy
           (drive:\path or drive:\path;drive:\path)

:: Logging Options :
/LOG:file :: Output to LOG file instead of autoverify.timestamp.log
/LOG+:file :: Appends to existing LOG file
/PI:n     :: Progress Indicator every n sec: default 10 (0 means no indicator)
/MAIL:receiver's e-mail address:SMTP Server:SMTP account:password:[port]
           :: (Example: /MAIL:m.muster@kost.ch:smtp.kost.ch:muster:12345)
/TESTMAIL :: Send a test e-mail and do nothing else
/FMB:n    :: Flush Mail Buffer after elapsed time: default 4 hours
/PMESS:n  :: Generate Performance Message after elapsed minutes: default non
/NMW      :: No Mail Warning in case of mismatch incidences
/U        :: Produce Verbose output, show each compared directory
/UU       :: Produce Very Verbose output, show each compared directory & file

:: Synchronize Options :
/COPY     :: Copy missing files and directories to destination with robocopy
/IPG:n    :: Inter-Packet Gap (ms), to free bandwidth on network: default 10
/PS:n    :: Packet Size for data transfer: default 64 kbyte
/NOC     :: No Orphan Control on destination file and directory
/NOD     :: No missing Directory processing
/NOF     :: No File verification process - simple directory listing

:: Retry Options :
/NMI:n   :: Number of Mismatch Incidences before shutdown: default 10
/RSF:file :: Output to Restart File instead of autoverify.rsf
/RST:n   :: Time Restart Information is valid in hrs: default 48
/R:n     :: Number of retries on failed read or copies: default 1 million
/W:n     :: Wait time between retries: default is 30 seconds
/LOOP    :: Loop until terminated with ctrl-c
/INTERVAL:hh:mm-hh:mm :: Loop only in dedicated time range: from-to

:: Help :
/UER     :: Version 1.05 of autoverify
/?      :: Display this help file
```

# Implementierung & Abnahme

- *autoverify* ist auf einem Server in der Staatskanzlei installiert und läuft dort nur nachts.
- Das Staatsarchiv hat inzwischen 98 GB (~16'000 Dateien) aus verschiedenen *Retrodigitalisierungs-Projekten* auf der Speicherbox abgelegt.
- Eine *dritte Speicherbox* wird in Erwägung gezogen.
- Die Übernahme von *originär digitalen Daten* ist noch in Planung.

