

PROTOCOLLO INFORMATICO PRESSO L'AMMINISTRAZIONE CENTRALE DELL'INFN REQUISITI TECNICI E NORME DI UTILIZZO PER LE APPARACCHIATURE DI ACQUISIZIONE DOCUMENTALE

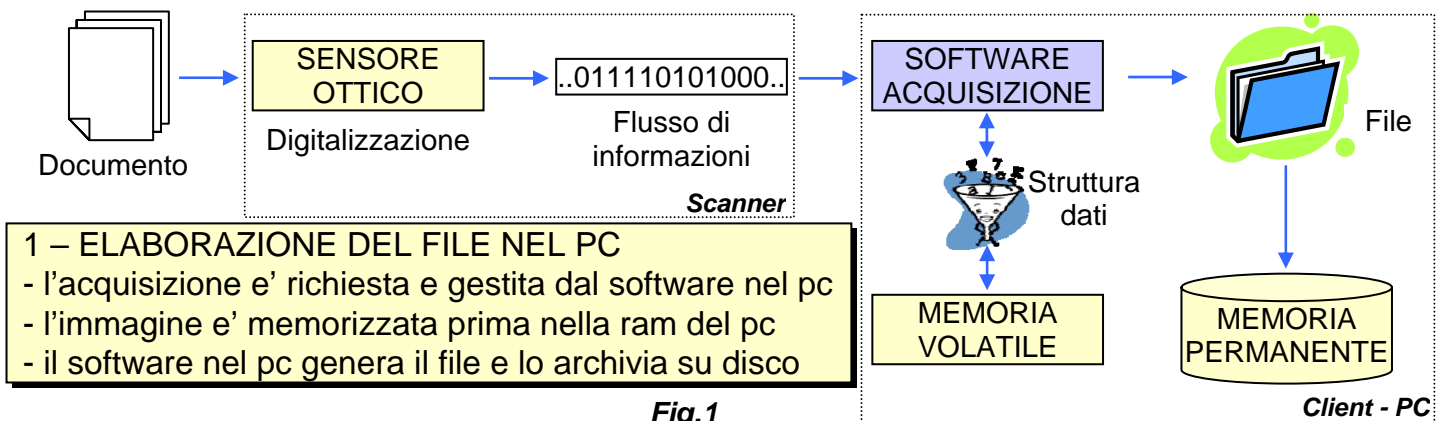
1. Meccaniche e metodologie di acquisizione dei documenti

L'acquisizione di un documento mediante scanner e' un processo mediante il quale i contenuti di alto livello, che compongono le pagine di un documento, sono codificati elettronicamente e memorizzati in un file.

Suddetto processo e' caratterizzato, in sequenza dalle due fasi distinte:

digitalizzazione: ogni pagina viene recepita dallo scanner come una 'scacchiera virtuale' composta da rettangoli: tanto piu' piccole essi sono, tanto piu' alta e' la sensibilita' dello scanner e la qualita' dell'output ottenuto. Per ogni rettangolo di questa scacchiera il sensore ottico legge le proprieta' cromatiche di uno solo dei vertici (*pixel*) codificandole in informazioni binarie. Questa fase restituisce quindi un flusso di informazioni elementari (*bit*) di dimensioni proporzionali al numero dei *pixel* che compongono la pagina e alle distinte sfumature di colore che sono distinte per ciascuno di essi.

acquisizione-elaborazione: il flusso di informazioni elementari converge in una opportuna struttura dati utilizzata per la codifica dell'immagine della pagina sia in memoria temporanea sia in memoria permanente sotto forma di file.



I processi di acquisizione possono essere classificati a seconda che la fase di elaborazione sia demandata al client (vedi **Fig.1**) o eseguita dallo stesso scanner (vedi **Fig.2**) .

L'elaborazione caratterizzata dall'uso indispensabile del client e' il metodo piu' semplice ed economico per acquisire documenti in formato elettronico: risulta tuttavia piu' dispendioso in termini di impatto sulle prestazioni del PC per il quale e' richiesto notevole quantita' di memoria disponibile, un processore veloce (per i PC si faccia riferimento agli Intel PIII o IV) e l'installazione di un'applicazione software complessa. Questo approccio e' quindi praticabile per volumi di produzione medio-bassi.

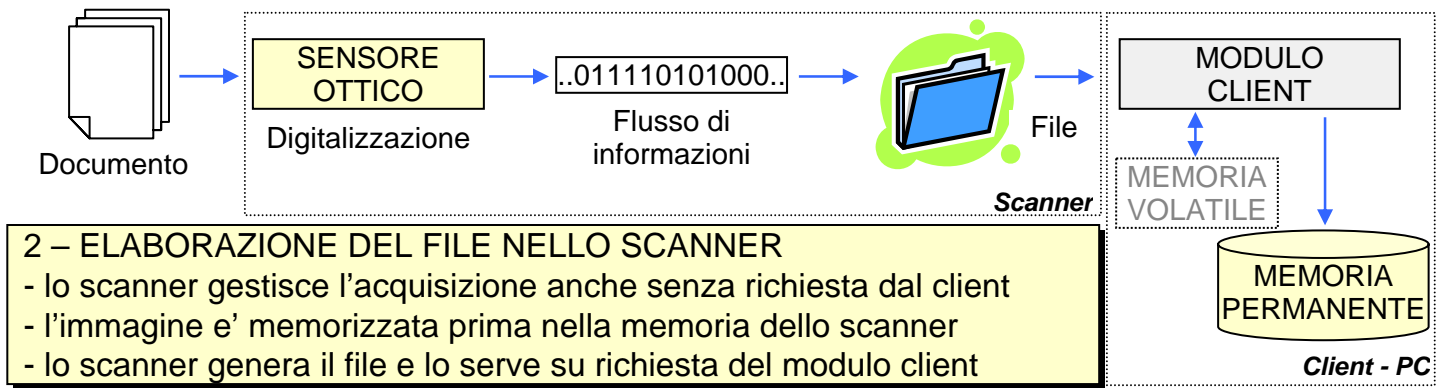


Fig.2

Affinche' l'elaborazione dell'immagine risieda nello scanner (vedi **Fig 2**), tale apparecchiatura dovra' essere piu' sofisticata, dovendo incorporare, oltre a proprie risorse di storage, anche la logica per la codifica e la generazione dei file: il costo dello scanner, mediamente molto piu' alto impatta in forma minore sulle prestazioni e sui requisiti dei client o PC.

2. Impatto sulle risorse

L'acquisizione elettronica di un documento richiede elevato impegno di memoria e risorse tra le quali :

- la classe/velocita' del processore;
- il *tranfert-rate* e/o la disponibilita' di banda nel collegamento fisico tra scanner e client (PC);
- le caratteristiche di elaborazione come la generazione di file PDF e il supporto per altri formati compressi.

A miglior comprensione della problematica in **Fig.3** e' fornita una tabella di comparazione delle dimensioni di un documento codificato nei differenti formati. Da notare che il risparmio derivante dall'utilizzo dei formati compressi (PDF, TIFF, JPEG, ecc) e' tale solo in termini di file di storage permanente che risulta doveroso in termini di efficienza di archiviazione degli allegati documentali.

IMPIEGO DI MEMORIA PER UNA PAGINA A4 NEI DIFFERENTI FORMATI DI FILE				
TESTO	GRAFICA 300dpi 256 liv. Grigio (8 bit)			GRAFICA 300dpi Colori (24 bit)
	Word DOC	TIFF non compr.	TIFF compr.	ACROBAT PDF
< 100 Kb	~ 8 Mb (~ 8.000%)	~ 6 Mb (~ 6.000%)	~ 1 Mb (~ 1.000%)	~ 25Mb (~ 25.000%)

Fig.3

3. Requisiti tecnici e parametri operativi

La tabella di **Fig.4** mostra i requisiti tecnici in base al carico medio giornaliero per postazione documentale con ulteriore discrezionalità sulla natura riservata dei contenuti dei flussi documentali.

Restano quindi individuati quattro profili funzionali i cui numeri a riferimento sono disposti in progressione crescente per fascia di costo.

Carico client (giornaliero)	REQUISITI TECNICI				NOTE OPERATIVE
	Fino a ~ 5-10 pg./doc. e poche unita' di documenti (~10)		Oltre 10 pg./doc. o decine di documenti		
Contenuto	Publicco	Riservato	Publicco	Riservato	
Sede di elaborazione del file di output	PC e Scanner		PC e Scanner		Tipo file: TIFF compresso PDF
Tipo di collegamento	Locale e Rete	Solo locale	Locale e Rete	Solo locale	Locale: SCSI- USB2 No Parallelo
Alimentatore automatico	SI + Un. fronte/retro	SI + Un. fronte/retro	SI + Un. fronte/retro	SI + Un. fronte/retro	
Risoluzione grafica	300-600 dpi	300-600 dpi	200-300 dpi	200-300 dpi	Norma: 200 dpi Risol. oriz.=vert.
Risoluzione cromatica	b/n scala grigi	b/n scala grigi	b/n scala grigi	b/n scala grigi	Non oltre 256 liv. grigio
Protocollo comunicazione	TWAIN e ISIS	TWAIN	TWAIN e ISIS	TWAIN e ISIS	
Memoria PC (Pentium IV sup.)	256- 512 MB	512 MB	256-512 MB	Oltre 512 Mb	
	2	1	4	3	

NB. gli eventuali valori consigliati per ciascun parametro sono indicati in grassetto Fig.4

In ogni caso, da un punto di vista operativo dovranno essere rispettate le seguenti specifiche:

- risoluzione grafica di acquisizione: **non superiore a 300 dpi** (consigliata 200 per caratteri alfanumerici di corpo superiore a 10 pt);
- risoluzione di acquisizione: **solo in scala di grigio con sensibilita' non superiore ai 256 toni**;
- formato per file di output : **preferibilmente PDF** o in alternativa TIFF compresso;
- protocollo di comunicazione postazione-scanner: **almeno TWAIN**.

Per le implementazioni che prevedono un collegamento locale diretto tra scanner e client risulta inoltre consigliato utilizzare dispositivi che prevedano:

- l'attivazione del processo di acquisizione direttamente da scanner, tramite azione su apposito pulsante per la scansione event-driven;
- la generazione automatizzata del file di output con memorizzazione diretta verso unita' di storage locali o di rete (Scan to drivers).

4. Norme di acquisizione e gestione degli allegati

Si consiglia di allegare preferibilmente documenti il cui formato nativo sia quello elettronico.

L'acquisizione degli allegati cartacei invece deve essere eseguita come attività indipendente dalla protocollazione per motivazioni principali :

funzionale: poiché è prioritario generare il record di protocollo dal quale ottenere il relativo numero e data, mentre l'acquisizione contestuale alla protocollazione potrebbe comportare latenze operative e ritardi sulla produzione;

tecnico: solo alcuni modelli sono supportati dall'applicativo di protocollo; inoltre errori nel processo di acquisizione contestuale potrebbe determinare il *fault* dei moduli software lato client del protocollo e impattare sulle funzionalità centrali dell'applicazione al punto che l'utente potrebbe non aver il corretto *feed-back* sull'esito della stessa protocollazione.

In ogni caso è opportuno allegare un solo documento alla volta di dimensioni possibilmente **non superiori ai 20 Mb**.

Fermo restando la preferenza per il formato PDF, in **Fig.5** viene mostrata una tabella riassuntiva dei tipi di file consentiti per ciascun formato di documento.

Formato documento	Applicazione di riferimento	Uso del corretto tipo di file per gli allegati		
		Consentito	Sconsigliato	Vietato
TESTO	Notepad Editpad	.TXT	-	-
	Wordpad	.RTF, .TXT	.DOC	-
	Ms Word	.RTF	.DOC, .DOT, .HTM, .HTML	-
TABELLE	Ms Excel	.TXT, CSV	.XLS, XLT,	-
GRAFICA	Vari	.TIF (compr.) .WMF, .CDR .PSD, .AI, EMF .DXF	.TIF (standard) .BPM, PCX, .JPG	-
PRESENTAZIONI	Ms Powerpoint	-	.PPT	-
CONTENITORI	Acrobat Writer	.PDF	-	.ZIP, .AVI
ALTRI FILE	-	-	-	.EXE, .DLL, .DRV, .VXD, .OCX, .VB, .VBA, .VBS, .JS, .ASP, APSX

N.B.: L'ESTENSIONE IN GRASSETTO CORRISPONDE AL FORMATO PROPRIO DELL'APPLICAZIONE

Fig.5

5. Implementazioni consigliate e modelli idonei

Dagli incontri intercorsi tra i Servizi AC-LNF riguardo alla gestione dell'archivio documentale e' emersa l'esigenza operativa di acquisire decine di documenti/giorno, ciascuno composto in media da un numero limitato di pagine.

In uno scenario di protocollo distribuito su postazioni documentali distinte, appare percorribile l'utilizzo di un'apparecchiatura di fascia media (**vedi profili 2 e 3 di Fig.4**) connessa generalmente in locale ad ogni postazione client di classe PC: per il protocollo di collegamento USB2 risulta quello consigliato.

Le postazioni client fisicamente vicine che protocollano volumi medio-alti di documenti non riservati (**vedi Fig.4**), possono condividere un unico scanner mediante l'utilizzo di *network storage in share*.

Nella seguente tabella sono illustrate le implementazioni consigliate indicando per ciascuna di esse un apparecchiatura scanner di riferimento.

REQUISITI E PARAMETRI OPERATIVI COMUNI	
Tipo di collegamento	Locale, USB2
Protocollo di comunicazione	TWAIN
Attivazione del processo di acquisizione	Automatizzata, direttamente da scanner
Sede di elaborazione file di output	Scanner o PC con scansione su disco diretta/non gestita
Caricamento fogli	Alimentatore automatico, unita duplex
Risoluzione grafica	200-300 dpi
Risoluzione cromatica	b/n o 256 livelli di grigio
Formato file di output	PDF
Caratteristiche salienti per il client	Postazione di classe PC Processore: Pentium IV o sup RAM: almeno 512 MB HD: 10 Gb disponibili per la scansione O.S.: MS. Windows XP Pro o sup. Software applicativo: MS Office 2003, Adobe Acrobat Professional v. 6.0 o sup.

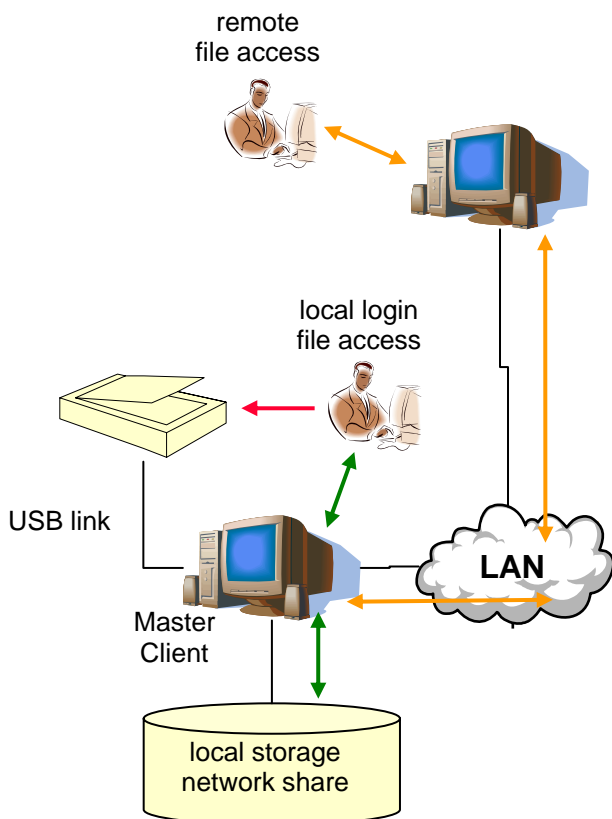
IMPLEMENTAZIONE CENTRALIZZATA

Scenario operativo consigliato per le attività degli uffici centrali come l'Ufficio Protocollo.

Un gruppo di client condivide lo stesso scanner collegato localmente ad uno di essi che costituisce la postazione master.

La postazione master esporta in rete un'area di storage come *network share* dove verranno depositati i file di output generati in forma automatizzata dal processo di acquisizione.

Devono essere previste opportune politiche di autorizzazione che garantiscano l'accesso ai suddetti file in favore dei soli account utente che effettuano il login sulla postazione master e/o remotamente su uno degli altri client.



Modello di riferimento:

Scanner **XEROX**, modello *DocuMate 262*

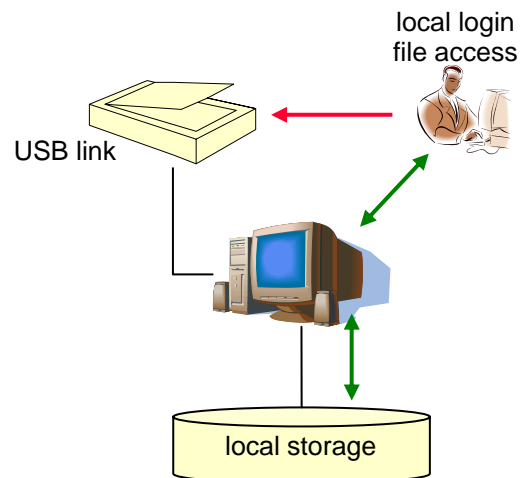
IMPLEMENTAZIONE DISTRIBUITA

Scenario operativo consigliato per le attività delle direzioni e servizi dell'A.C.

Ad ogni client è collegato localmente uno scanner.

L'acquisizione genera in forma automatizzata un file di output che è memorizzato nelle storage locale del client.

Solo gli utenti autorizzati al login interattivo sul client accedono ai suddetti file e li allegano al Protocollo.



Modello di riferimento:

Scanner **XEROX**, modello *DocuMate 152*

Frascati, 12 dicembre 2006

Nunzio Amanzi