



# Versione "Hot Work In Progress"

- Introduzione del concetto di profilo. Abolizione del file xusers.xml a favore di un diverso sistema di autenticazione ed adozione di profilo di sistema e d'archivio.
- Modalità Blind, con .conf.xml minimale.
- Impostazioni di indicizzazione direttamente nel campo che viene sottoposto al server.
- Revisione generale delle politiche di *locking/flushing* ed uso della cache del BTree in molteplici casi.
- Adozione del B+Tree in tutti i punti possibili, sia nelle attività che riguardano il *file system* sia per le operazioni di indicizzazione in memoria.

Data	Mod	Ver		Sintomo o Segnalazione
16/09	xw	23.0.0.1_hot_wip "EVO" 22.1.3.10	NC	Search & Replace incapace di compiere un'attività precisa.
<p>Le procedure di <i>Search &amp; Replace</i> risultano incapaci di compiere un'attività più precisa e capillare. Nella fattispecie è incapace di saltare i documenti in cui la modifica richiesta non abbia senso e non venga applicata ne di indicare, nei casi di errore, quali documenti siano interessati dagli errori ed in che modo</p> <p>Implementato con l'aiuto di Gregorio un nuovo procedimento in grado di identificare le attività positive da quelle non effettuate e regolarsi di conseguenza.            Dettagli nella <a href="#">documentazione del comando</a>.</p>				
10/09	xw	23.0.0.1_hot_wip "EVO" 22.1.3.9	CE	Crash del server in accesso al vocabolario.
<p>A quanto è possibile vedere, il buffer globale destinato al trattamento delle chiavi era sottodimensionato rispetto alla massima dimensione lecita per una chiave. Se pure è raro che essa si presenti negli indici dei documenti, è invece facile che si manifesti nel thesaurus che prevede chiavi molto lunghe e composte da due parti entrambe ampie.            Vds. GWT000855.</p>				
03/09	xw	23.0.0.1_hot_wip "EVO" 22.1.3.9	CE	L'esecuzione di un <i>Search &amp; Replace</i> da esiti errati.
<p>Pur eseguendo correttamente il proprio compito, l'attività di <i>Search &amp; Replace</i> indicava erroneamente quante e quali fossero i salvataggi effettuati e con quale esito..</p>				
09/08	xw	23.0.0.0_hot_wip "EVO"	CA	<a href="#">Emissione Candidate da testare internamente</a>
02/08	xw	23.0.0.0_hot_wip "EVO" 22.1.3.7	CE	L'import da <i>WatchDoc</i> di documenti con encoding non corrispondente a quello dichiarato nel file presenta una condizione d'apparente correttezza ma lascia il file .fail sul proprio percorso.
<p>L'intervento effettuato nella precedente 22.1.3.6 non azzerava il codice d'errore vanificando quindi l'operazione di salvataggio.</p>				
30/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.7	CE	La costituzione del catalogo che preveda l'applicazione di un foglio di stile XSLT alle volte, pur non dando errore, non produce alcun risultato.
<p>Il difetto va ricercato nel fatto che la funzione riceve puntatore e dimensione del buffer da elaborare ma se vi deve applicare anche il <i>namespace</i> xw: per poter interpretare correttamente gli allegati, compone mane il buffer da elaborare perché ignora la size notificata. In inserimento e modifica la cosa non si avverte ma nel fare catalogo da file (quindi anche da <i>WatchDoc</i>) l'errore può presentarsi.            Rettificato l'uso della dimensione anche in quel punto e corretto l'errore.</p>				
27/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.7	CE	La combinazione di <i>ranges</i> in forma di "Or List" ed altre estensioni (genere e numero, thesaurus, somiglianza, fonetica, etc.) non da l'esito atteso. Le estensioni non risultano applicate.
<p>L'errore è dovuto al fatto che il <i>range</i> in forma di "Or List" veniva trattato come tutti i ranges, ovvero estraendo esattamente quei termini dal vocabolario, senza tenere in considerazione la possibilità che da essi si derivassero altri termini. Tutte le altre forme di <i>range</i> si avvalgono <u>espressamente ed intenzionalmente</u> dei termini presenti nel vocabolario, e solo di essi, ma questo <i>range</i> fa naturalmente razza a se.            Vds. eGW#000840.</p>				
21/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.7	CE	La presenza di parentesi quadre nella frase di ricerca che non intendano racchiudere un campo conduce a comportamenti inesatti/inattesi. La forma [campo1]=[campo2] non da errore ma non da esito..
<p>La richiesta (Vds. eGW#000770) non può essere accolta in quanto un tentativo di abbattere l'ambiguità data dall'uso improprio di parentesi quadre al fine di considerarle alla stregua di una strina qualsiasi impatta con la ricerca con <i>Anyalias</i> e la ricerca multi archivio, oltre a confliggere con alcuni casi di univocità.            La correzione si limita ad inibire e considerare sintatticamente errata la forma [campo1]=[campo2].</p>				
19/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.7	CE	Le ricerche per numero documento confliggono con le estensioni, ad esempio, per genere e numero.



Data	Mod	Ver		Sintomo o Segnalazione
				Difetto riscontrato e corretto. Vds eGW#000836.
19/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.7	CE	Gli indici corrotti dilagano a macchia d'olio e bloccano il funzionamento dell'archivio.
				La problematica era già stata trattata in precedenza, per interrompere l'uso della catena dei liberati di un <i>Buddy File</i> quando essa risultasse corrotta. La correzione era però insufficiente in quanto utilizzava comunque una prima area già riutilizzata con riflessi molto pericolosi sulla consistenza di quell'indice. L'attuale correzione (Vds. eGW#827) è pensata per evitare anche questo problema.
14/07	xw	23.0.0.0_hot_wip "EVO"	NC	Nel visualizzatore eventi appaiono una serie di segnali che risultano errori pur non essendo nulla di serio e creano un ingiustificato allarme.
				Compiute due verifiche. In primo luogo non vengono più registrate come errori i mancati logs che si cerca di inviare a xwls prima ancora che l'interfaccia per fare logging sia stata del tutto avviata. In secondo luogo, tutte le informazioni inviate al <i>Log Giornale</i> vengono riversate anche sul log di sistema <sup>1)</sup> quindi si evita di indicare come <i>Warning</i> segnalazioni fisiologiche o innocenti. (Vds. segnalazione eGW000744).
13/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.7	NC	Evitare gli indici degli allegati quando essi si presentino di identico contenuto.
				La funzionalità è stata concepita e sviluppata per avvantaggiarsi in scenari in cui ci siano allegati identici (Vds. segnalazione eGW000735) cosa che avviene non di rado ad esempio quando si allegano più versioni di un determinato documento in cui il testo non risulti di fatto cambiato. Malauguratamente, per la suddetta segnalazione, questo intervento risulta inutile.
12/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.7	CE	Le attività massive che interessano conversioni XSLT sono estremamente lente.
				Tentata un'ottimizzazione mantenendo caricato ed utilizzabile l'ultimo foglio di stile XSLT utilizzato al fine di evitare un inutile overhead nel ricaricare continuamente un oggetto sempre identico. Purtroppo il guadagno in termini prestazionali non è ancora sensibile come speravamo, perché limitato all'XSLT e non ha effetti sul tempo necessario al caricamento DOM del XML da elaborare.
01/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.6	CE	Gli errori di encoding rendono inusabili le vecchie applicazioni.
				Il server non deve accettare che documenti dichiarati in encoding iso8859-1 contengano, in realtà, caratteri appartenenti all'encoding windows-1252. La modifica di recente introduzione che inibisce il salvataggio di simili documenti, pur essendo corretta, causa diversi problemi alle applicazioni esistenti. Per ovviare a questo si può introdurre, nei vecchi archivi, una voce di profilo che rende più lasco il controllo. Il server si limiterà ad indicare che c'è un errore di encoding senza provocare un errore bloccante. <profile type="arc.test_encoding" value="false"/>
01/07	xw	23.0.0.0_hot_wip "EVO" 22.1.3.6	CE	Esprimendo un range di valori sotto forma di <i>or list</i> , il programma ha un comportamento strano se uno di tali valori è il valor vuoto.
				Ci sono due ordini di problemi. Se il valor vuoto viene espresso con la sequenza &null; la presenza del punto e virgola confonde l'interprete dei ranges in quanto anch'esso è uno dei separatori di range. Se il valor vuoto viene espresso con "&null;" il comportamento è corretto mentre se si indica semplicemente "" l'esito non comprende il valor vuoto. Sul primo problema c'è poco da fare, la sintassi rischia la confusione quindi si deve fare attenzione nell'utilizzo, il caso della sequenza "" è stata equiparata a quella del termine "&null;". Ammissa come valida anche la forma [campo]={valore} ovvero [campo]={ "valore" } che viene trattata come <i>or list</i> anche se l'elenco è composto da un solo valore. Vds eGroupWare #820.
29/06	xw	23.0.0.0_hot_wip "EVO" 22.1.3.5	CE	La ricostruzione integrale di un archivio, in presenza di difetti che ne comportino la rimozione del file .stat.xml, causa la perdita dei seriali.



Data	Mod	Ver		Sintomo o Segnalazione
				<p>Il server assume che se un archivio esiste i suoi seriali si conservano, se non esiste si butta ogni cosa si trovi lungo la strada, assumendo che siano dati errati. Per sapere se l'archivio esiste ci si basa sul file <i>nomearchivio.stat.xml</i> in assenza del quale si assume che ci sia di buono solo il <i>.conf.xml</i>.</p> <p>In questo caso si perde il <i>.ser.xml</i> in quelle occasioni in cui sia stato necessario cancellare alcuni file dell'archivio perché danneggiati o incompatibili con la versione di server che si va ad usare.</p> <p>Visto che il flag di salvaguardia dei seriali esiste dalla notte dei tempi ma mai nessun client ne ha fatto uso, il server ora agisce d'ufficio salvaguardando sempre i seriali ed i thesaurus ed ignorando l'eventuale flag ricevuto.</p> <p>IL server 23 mantiene però la possibilità di gestire la rimozione dei seriali in caso di comando di "drop" d'archivio per gli archivi Lucene, ovvero gli archivi "blind".</p> <p>Vds eGroupWare #788.</p>
29/06	xw	23.0.0.0_hot_wip "EVO" 22.1.3.5	CE	Titoly lazy di più archivi si rubano continuamente la scena, rendendo molto più lento del necessario il compito di ricostruzione lazy dei titoli.
				<p>Il problema dipende dal fatto che mente si sta facendo un'operazione di lungo corso, come ad esempio il rifacimento dei titoli, il server tenta di fare anche altre attività eventualmente prioritarie. Per tale ragione da un po' di tempo a ciascuno dei compiti collocati nella directory lazy contando che, chi più chi meno, arriveranno tutti a termine senza particolare ritardo. Dalla versione corrente del server i compiti lazy sono stati razionalizzati e quindi è stata data priorità 'A' alle indicizzazioni off-line, priorità 'B' alle indicizzazioni delle modifiche dei documenti e priorità 'Z' ai titoli. In questo modo, mente si stanno facendo i titoli di un archivio, altri titoli non distoglieranno il primo dal suo compito mente attività a priorità superiore verranno svolte comunque prima interrompendo il flusso di attività in corso.</p> <p><u>Inoltre da adesso i titoli off-line vengono sempre fatti lazy, salvo diversa disposizione del chiamante.</u></p> <p>Vds eGroupWare #796.</p>
09/06	xw	23.0.0.0_hot_wip "EVO" 22.1.3.4	CE	L'indicizzazione dell'archivio docway di Asolo da crash del server.
				<p>Il problema è dovuto al fatto che si indicizza in un certo "lotto" una chiave che appare due volte. Ciò significa che è stata inserita due volte nel BTree in memoria considerandole distinte ma che risulta la stessa chiave nel BTree su disco. Questo comporta che la chiave venga scaricata due volte producendo un file .LOT inconsistente (non può essere scandito in modo corretto) che provoca un crash in fase di scarico finale.</p> <p>Corretto l'errore intervenendo sulla funzione di normalizzazione delle chiavi.</p>
28/05	xw	23.0.0.0_hot_wip "EVO" 22.1.3.4	NF	Semplificare la procedura per la realizzazione di archivi per CD.
				Realizzato un singolo comando in grado di compiere tutti i passi necessari alla realizzazione di un archivio che sia il clone integrale o semplicemente di una parte di un altro archivio. L'archivio così ottenuto può essere già pronto per l'uso su un supporto ottico e con una distribuzione di server in grado agire su archivi così protetti.
26/05	xw	23.0.0.0_hot_wip "EVO" 22.1.3.4	CE	Non si riesce a compiere un'importazione via WatchDoc su alcuni archivi.
				<p>Se un archivio privo di <i>file_location rule</i> la modalità usata dal server per identificare in quale file XML scrivere il documento è quella detta <i>single</i><sup>2)</sup>. La procedura che determina il nome di tale file è la stessa usata per gli allegati. In tale frangente, essendo fondamentale l'univocità, un archivio non completamente indicizzato causa errore. Questo è lo scenario che si presenta dopo l'importazione del primo documento via <i>WatchDoc</i>.</p> <p>Allentata la stretta del controllo sull'univocità quando la procedura viene utilizzata per scopi legati ai files XML anziché agli allegati.</p>
10/03	xw	23.0.0.0_hot_wip "EVO"	CE	L'indicizzazione di alcuni allegati manda in crash il server.
				<p>Il problema va ricercato nel fatto che l'adozione dell'OCR Tesseract produce alle volte documenti contenenti una grande quantità di dati errati il cui formato, di fatto, è UTF-8 anche se il file è del tutto testuale. Dal server di classe 22 si introduce un test che fa sì che si proceda alla conversione del contenuto secondo il detto encoding così da compiere un'indicizzazione dei contenuti il più corretto possibile.</p> <p>Malauguratamente quest'attività può condurre ad una crescita anche sensibile della stringa che rappresenta il contenuto dell'allegato in quanto da UTF-8 a WinLatin1<sup>3)</sup> molti caratteri possono risultare non traducibili e vengono quindi traslati in <i>entity</i> numeriche. La funzione a ciò preposta allocava <i>un po' di spazio in più</i> senza valutare correttamente il dafarsi.</p> <p>Nel server 22 si è provveduto a compiere un'allocazione di memoria più vasta, nel 23 si passa all'uso di <code>std::string</code> così da aggirare completamente l'ostacolo.</p> <p>Vds eGroupWare #750.</p>
04/03	xw	23.0.0.0_hot_wip "EVO"	CE	A seguito di un'azzeramento d'archivio il risultato dello stesso è inutilizzabile.
				<p>Mancando una reale condizione di test sulla concorrenza tra l'azzeramento di un archivio, considerato abbastanza definitivo per non dover prendere particolari precauzioni, ed eventuali operazioni di inserimento e modifica, è possibile, in linea teorica (ed anche in linea pratica a giudicare dalla segnalazione eGW #747) che un server sovrascriva il file <i>nomearchivio.stat.xml</i> e le header degli altri files dopo che un altro server ha provveduto all'azzeramento.</p> <p>L'archivio risultante se tutto va bene, non può più essere aperto<sup>4)</sup> ne quindi si riesce a ricostruirlo da capo se non rimuovendo a mano i files.</p> <p>Si da il caso, però, che il server che l'ha azzerato, convinto della sua validità sulla base di quello che ha in <i>cache</i>, consente operazioni come la costituzione della mappa e dei titoli con risultati drammatici<sup>5)</sup>.</p> <p>Compiuto un intervento che si assicura di bloccare l'archivio esistente prima di rimuoverlo così da assicurarsi che l'archivio risultante sia corretto.</p>



Data	Mod	Ver		Sintomo o Segnalazione
04/03	xw	23.0.0.0_hot_wip "EVO"	CE	L'accesso al vocabolario con pattern causa condizioni di <i>loop</i> .
				La verifica della presenza di un pattern comune in caso di ricerca del chiave greater non teneva conto correttamente, nelle successive letture 'next', che la radice fosse per lo meno comune. Vds. <a href="#">GWT000639</a>
04/03	xw	23.0.0.0_hot_wip "EVO"	CE	In fase di spegnimento servizi capita che un server slave rimanga attivo e consumi CPU.
				L'errore si ha quando un server slave riceve il comando che lo informa che deve chiudersi. Tale comando, asincrono, viene gestito da un apposito <i>handler</i> il quale è pensato per essere eseguito in modo recursivo ogni volta che esso rileva che c'è un comando ancora pendente. Il bit del comando di chiusura non viene abbattuto in quanto è utile sapere che la chiusura è stata richiesta. Inibita la reiterazione qualora il solo comando pendente sia quello di chiusura. Vds. <a href="#">GWT000559</a>
04/03	xw	23.0.0.0_hot_wip "EVO"	CE	L'esecuzione massiva della ricostruzione titoli produce un file di titoli inutilizzabile e quindi causa sensibili rallentamenti a tutto il sistema.
				Il problema s'è manifestato da quando, nella costituzione dei titoli, si è deciso di procedere dall'ultimo al primo <sup>6)</sup> anziché in ordine numerico naturale. La procedura che provvedeva a sbiancare la parte di file <i>.tip</i> non ancora valorizzata compiva un errore banale sovrascrivendo anche la header con risultati disastrosi. Vds. <a href="#">GWT000636</a>
03/03	xw	23.0.0.0_hot_wip "EVO"	CE	L'applicazione di un filtro nell'accesso al vocabolario che presenti un asterisco come primo carattere produce una stringa vuota.
				Il confronto sul pattern era corretto ma, malauguratamente, tale confronto non doveva limitarsi ad affermare se la stringa fosse corrispondente o meno, ma anche se si fosse ancora in un <i>range</i> utile per evitare che il controllo sul pattern scandisca inutilmente tutto il vocabolario. Per verificare se si rientra nel <i>range</i> si controlla la corrispondenza della parte <i>radice</i> della chiave che in presenza di <i>Wild Card</i> come primo carattere non dava un risultato significativo. In questo caso, ora, si asserisce di essere sempre nel <i>range</i> .
19/02	xw	23.0.0.0_hot_wip "EVO"	CE	Indici corrotti.
				Problema dipendente dalla presenza di un carattere 'p' che veniva maldestramente interpretato dalla funzione di normalizzazione. Il problema è stato arginato mettendo una vera e propria pezza nelle funzioni che inseriscono le chiavi nel BTree/CBtree/BTreeMem ed imponendo un ulteriore filtro nella funzione di indicizzazione incrementale o differenziale che inserisce le chiavi visto che in alcuni casi, meno rari del previsto, si inseriva una chiave vuota. Vds <a href="#">GWT000596</a> .
28/01	xw	23.0.0.0_hot_wip "EVO"	CE	Le attività di log impattano sulle performance più di quanto si attenda.
				Cambiata la politica di logging usando un thread parallelo <i>lockless</i> che compie il logging mentre il thread principale procede nelle sue attività. Se lo sviluppo si dimostrerà efficace a simili threads verrà dato il compito di fare altre operazioni, come la registrazione delle attività <i>lazy</i> o del <i>registro</i> ..
27/01	xw	23.0.0.0_hot_wip "EVO"	CE	Caccia alle performance.
				Cambiata la politica di apertura degli archivi per riconoscere i comandi <i>Archiveless</i> senza dispendio di tempo. Abolite inutili memset di aree di memoria quando non necessario. Ridotta l'allocazione di memoria che viene fatta per fare <i>cache</i> della selezione in corso di esecuzione che era eccessiva e spostata sullo stack.
22/01	xw	23.0.0.0_hot_wip "EVO"	CE	In assenza del file <i>context.stat.xml</i> esso viene creato alla prima registrazione ma viene realizzato errato.
				Corretta la stesura del file nel quale, se assente, un attributo <i>id</i> veniva realizzato erroneamente come elemento.
30/12	xw	23.0.0.0_hot_wip	CE	Compire <i>BenchMark</i> non fornisce facilmente indicazioni statistiche utili.
				Introdotta una modalità di stesura di un log di <i>BenchMark</i> che semplifica la vita al tester. Vds <a href="#">GWA001344</a> .
29/12	xw	23.0.0.0_hot_wip	CE	Nel compiere dei <i>BenchMark</i> con altre versioni del server verificata una lentezza in modifica documenti.
				Il problema è stato arginato adottando l'uso della <i>cache</i> del BTree anche in modifica, cosa che non era mai stata fatta. Vds <a href="#">GWA001344</a> .
24/12	xw	23.0.0.0_hot_wip	NF	Nel compiere dei <i>BenchMark</i> con altre versioni del server nota che viene compiuto un numero irragionevole di letture, seek, scritture e flush dei files dell'archivio.



Data	Mod	Ver		Sintomo o Segnalazione
				La gestione delle letture delle header, delle loro scritture e dei flush degli archivi era <i>sparsa</i> un po' ovunque e costringeva a rileggere continuamente per esser certi di avere a disposizione una versione affidabile. L'intera gestione degli accessi alle header è stata riveduta e migrata, per competenza, al più basso livello disponibile, ovvero al livello del singolo file. Ora, quando si blocca un archivio e si provvede al caricamento delle header dei files, gli stessi vengono etichettati in modo da evitare ogni ulteriore reload delle header non necessario. In questo modo viene riveduto anche tutto il comportamento delle librerie dinamiche tornando ad avere un comportamento finalmente coerente. La verifica sul campo dimostra che il numero di letture e scritture è sensibilmente diminuito con evidenti benefici.
17/12	xw	23.0.0.0_hot_wip	NF	Nello sviluppo della sintassi per Lucene ci sono alcune forme di range non previste.
				I range mancanti sono quelli che potremmo definire misti. ExtraWay server prevede sia i ranges inclusivi che esclusivi, ovvero i ranges nei quali gli estremi indicati debbano appartenere o meno all'esito finale della selezione. Lucene prevede, oltre ad essi, i ranges che risultano inclusivi solo per uno dei due estremi. Per integrare questo comportamento sono stati aggiunti i modificatori dei ranges '>' e '<' che si aggiungono a quelli già indicati L'espressione: [campo]={minval>maxval}  Indica che il range richiesto dovrà <u>comprendere</u> l'estremo espresso con <i>minval</i> ed <u>escludere</u> l'estremo espresso con <i>maxval</i> Analogamente l'espressione: [campo]={minval<maxval}  Indica che il range richiesto dovrà <u>escludere</u> l'estremo espresso con <i>minval</i> ed <u>includere</u> l'estremo espresso con <i>maxval</i> .
03/12	xw	23.0.0.0_hot_wip 21.2.1.20	CE	I timestamp interni, senza il carattere 'T' del formato ISO, non vengono trattati adeguatamente.
				Esteso il trattamento delle chiavi che rappresentano un timestamp <sup>7)</sup> anche in assenza del separatore 'T' che ne qualifica la forma riconosciuta ISO.
30/11	xw	23.0.0.0_hot_wip	NF	Le selezioni in modalità Lucene richiedono che il grado di somiglianza sia configurabile per singola ricerca.
				eXtraWay Server prevedeva esclusivamente l'indicazione della ricerca per somiglianza dei termini applicata all'intera selezione. Dalla versione in esame è stata introdotta la possibilità di indicare invece una somiglianza per singolo termine o, più precisamente, per il singolo campo sul quale si sta compiendo la selezione stessa. Vedasi la <a href="#">documentazione relativa</a> .
26/11	xw	23.0.0.0_hot_wip	NF	Per compiere diverse tipologie di selezione, tra le quali quelle in stile <i>Lucene</i> , sono state implementate le operazioni di ricerca in cascata.
				Vedasi la <a href="#">documentazione relativa</a> .
25/11	xw	23.0.0.0_hot_wip	NF	Per realizzare archivi da utilizzarsi nello scenario <i>Lucene</i> si devono costituire archivio con caratteristiche particolari.
				Introdotta il concetto di Archivio "Blind" che ciecamente accetta documenti ed alias di ricerca anche in assenza di una specifica configurazione. Vedasi la <a href="#">documentazione relativa</a> .
12/11	xw	23.0.0.0_hot_wip	NF	Il compattamento archivio viene fatto anche se inutile.
				Con l'adozione del B+Tree cambia completamente la filosofia del compattamento. Il compattamento del file .idx.bpt può risultare molto utile (la procedura che lo genera lo riempie lasciando molto sfrido) e viene eseguita a parte rispetto al resto. Il compattamento di VCB e REF viene ora effettuato solo se nel Ref c'è uno spazio libero superiore ai 256Mb o in generale se lo spazio libero supera il 10% dello spazio totale del file REF. Questo in particolar modo perché a seguito di indicizzazione <i>Off-Line</i> il ref è praticamente già compattato. L'unica caratteristica che manca ai 3 files è quella di essere in ordine (l'ordine dei record del VCB corrispondente all'ordine delle chiavi nel B+Tree e conseguentemente corrispondente all'ordine dei blocchi del REF) cosa che aiuterebbe in accesso sequenziale (ricerche per range, accesso ai vocabolari e simili) ed eventualmente per le applicazioni distribuite su supporto ottico.
11/11	xw	23.0.0.0_hot_wip	NF	Dover affrontare archivi di dimensioni sempre maggiori impone prestazioni più imponenti.
				Compiuti considerevoli interventi nell'indicizzazione <i>Off-Line</i> che portano archivi di dimensioni considerevoli ad essere indicizzati in tempi molto più rapidi. <small>Per portare un esempio, l'indicizzazione dell'archivio Agi che richiede, per una versione 22.1.2.0 un tempo di 8h07' viene compiuta dalla versione attuale in 4h59'. Un archivio campione di Regione Veneto è passato da 13h05' a 9h22'.</small> Il programma fa attualmente un uso maggiore della Ram, utilizzando quindi una cache più <i>spinta</i> . A tale scopo è stato modificato il file di configurazione <sup>9)</sup> nella sezione [hs] aggiungendo una nuova voce CacheSize il cui valore identificato come segue:
				<b>CacheSize presente e diverso da '0'</b> Rappresenta il numero di <b>Mb</b> da usare per la Cache. Ora si utilizzano due distinte forme di Cache quindi metà dello spazio indicato verrà usata per la Cache del B+Tree e l'altra metà per la lavorazione delle catene dei riferimenti. I valori di NodiCacheMax e KRAM vengono ignorati.



Data	Mod	Ver		Sintomo o Segnalazione
				Il valore della Cache viene ottenuto prendendo in considerazione il valore di NodiCacheMax <sup>10)</sup> e KRAM <sup>11)</sup> . Entrambe i valori vengono tradotti in <b>Mb</b> e CacheSize assume il doppio <sup>12)</sup> del massimo dei due valori.
				Vengono inoltre aboliti i seguenti valori di configurazione <sup>13)</sup> : AreaDiLavoro, AreaDiLavoroMax, FiltroInizio e FiltroFine mentre, come già detto, i valori di NodiCacheMax e KRAM vengono presi in esame solo in assenza di CacheSize. Il valore di NodiCache cambia il proprio valore di default da 510 a 1020.  <b>Nota:</b> Si consideri che con server sino alla classe '22' una configurazione tipo poteva prevedere un valore di NodiCacheMax pari a 192000 e di KRAM pari ad 80000. In questo scenario si finiscono con l'utilizzare 96Mb di Ram per la Cache del BTree e poco meno di 80Mb per il trattamento dei lotti. Questi due ammontare di memoria venivano allocati e liberati in momenti diversi quindi l'occupazione effettiva massima corrispondeva a 96Mb. Dal server 23 si sfrutta, in ognuna delle due fasi dell'indicizzazione <i>Off-Line</i> il doppio della Ram il che comporta un utilizzo medio di 192Mb di Ram. Le impostazioni suddette equivalgono quindi ad indicare esclusivamente CacheSize=192.
30/10	xw	23.0.0.0_hot_wip	NF	La gestione di un archivio richiede, anzi impone in alcuni casi, che per ottenere i risultati voluti in termini di indicizzazione sia necessario modificare il file di configurazione d'archivio per introdurre nuovi tipi di canali.
				Introdotta il concetto di canali autodefiniti. Rimane sempre vero quanto detto per attributi ed elementi che vengono indicizzati reciprocamente come monovalore e multivalore, alfanumerici e con l'applicazione della StopList <sup>14)</sup> solo per gli elementi. Se un elemento <sup>15)</sup> è corredato da alcuni attributi appartenenti al <i>namespace</i> <i>xw:</i> , essi vengono interpretati come modificatori dell'espressione della chiave e producono una chiave che ha particolari caratteristiche, eventualmente distinte dalle condizioni di default. Sono ammessi alcuni valori noti che derivano direttamente da quelli già esistenti nel file di configurazione d'archivio ed altri di <b>nuova concezione</b> .
				<b>Attributo</b>
				<b>Valori Ammessi (in Verde i valori di default, in Grigio i valori supportati ma di più raro utilizzo)</b>
				<b>xw:key_style</b>
				single, one, multi, double, skip, one-md5, one-md5cs, double-md5, double-md5cs, spatial.
				<b>xw:value_type</b>
				date, num, alfa, file, alnum.
				<b>xw:instance</b>
				Qualsiasi valore positivo o <b>negativo</b> (es.: true, yes, on, 1 ovvero false, no, off, 0)
				<b>xw:store</b>
				Qualsiasi valore <b>positivo</b> o negativo (es.: true, yes, on, 1 ovvero false, no, off, 0). Indica se si richiede o meno la conservazione del dato dopo aver provveduto a farne gli indici. <a href="#">All'atto della stesura della presente documentazione, quest'opzione viene prevista ma non implementata: il server farà comunque storage del dato.</a>
				<b>xw:analyzer</b>
				Identificatore univoco dell'analizzatore del testo da applicare sul contenuto del canale. <a href="#">All'atto della stesura della presente documentazione, quest'opzione viene prevista ma non implementata: il server applicherà comunque l'analizzatore/tokenizzatore di default.</a>
21/10	xw	23.0.0.0_hot_wip	CE	Il test di <i>Wellformedness</i> dei documenti può essere inibito con conseguenze anche gravi sui contenuti degli archivi.
				Ignorata l'inibizione introdotta in versioni molto vecchie. Ora il test avviene sempre e comunque nelle fasi di salvataggio di singoli documenti come di serie di documenti <sup>16)</sup> . Rimane la possibilità di inibire il test in fase di catalogo, assumendo che tutti gli XML dell'archivio, inseriti compiendo il test, siano intrinsecamente validi.
03/08	xw	23.0.0.0_hot_wip	NF	Le modifiche registrate ne file di registro mostrano il documento com'era e non com'è. Dei documenti meramente inseriti, in sostanza, non v'è traccia alcuna.
				Si deve compiere un'attività più precisa, registrando i documenti inseriti e, dove presenti, sia l'originale che la modifica se pure con una crescita del registro potenzialmente sensibile.
??/08	xw	23.0.0.0_hot_wip	NF	Rifacendo il catalogo di un archivio, le date di inserimento e modifica degli stessi evidenziate nella risposta al caricamento ovvero disponibili sotto forma di vocabolari subiscono una trasformazione ed appiattimento.
				Il vocabolario delle date di inserimento e modifica dei documenti viene ora prodotto sulla base delle <i>processing instruction</i> presenti nei documenti stessi divenendo indipendente dalla ricostruzione del catalogo.

1)

Grazie ad una configurazione del file *xwls.conf.xml*

2)

Che corrisponde ad un singolo file, numerato progressivamente, per ciascun documento



3)

L'encoding usato per *default* nel fare gli indici

4)

La header del file `.idx` manda ad una root che non esiste assolutamente

5)

Lo `.stat.xml` cita documenti che non ci sono, appare un catalogo che conta il doppio dei documenti e la prima metà non esiste

6)

In questo modo si privilegiano i titoli dei documenti più recenti che sono maggiormente candidati ad essere consultati

7)

Vale a dire una stringa che indica anno, mese, giorno, ore, minuti e secondi

8)

In questo caso, però, si paragona un server compilato *debug* ed uno compilato *release*

9)

`xw.ini`

10)

che rappresenta il numero di nodi di 512 bytes

11)

che rappresenta un valore espresso in **Kb**

12)

per le ragioni indicate nel punto precedente

13)

Che quindi vengono ignorati qualora presenti

14)

ovviamente in sola ricerca

15)

questa soluzione non si applica agli attributi

16)

ad esempio via WatchDoc