Ceralon Acumen
Performance summary
for multi-server and single server
environments on HP DL 380

Acumen is an integrated e-Discovery solution that delivers import, review, production and export functionality on a scalable multi-tiered server architecture. Ceralon Acumen provides a familiar browsing interface that encapsulates robust security, powerful searching and culling, adaptive workflow, native document viewing, and production capabilities.
The following tests, performed at HP's PTAC, demonstrate Acumen's performance on an enterprise environment.

Client configuration options

Clients may access Acumen modules by either connecting to them on a centrally hosted application server via Terminal Services or Citrix or by installing them directly on their desktops:

Terminal Services or Citrix
Acumen application modules (Import, Review, Export, and System Admin) are installed on the application tier of the 3-tier server environment. Users connect to the server hosting the Acumen application modules using Citrix or Terminal Services.
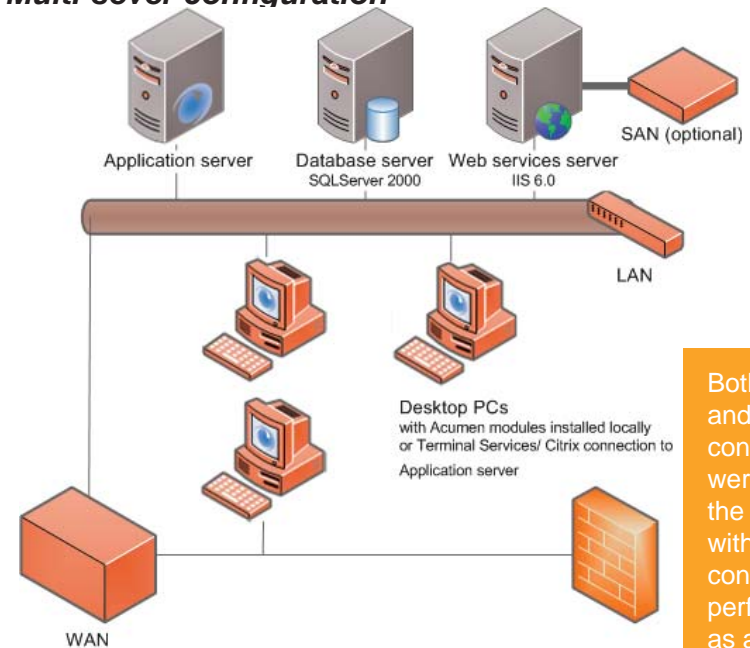
Desktop clients
Users install Acumen modules for Review, Export, and System Admin on their multi-purpose desktop PC. The Acumen import module is installed on one or more dedicated workstations. (The import module requires a dedicated workstation due to heavy memory utilization and Outlook PST Processing restrictions) The web tier and database tier of Acumen are hosted on centrally located servers.

Users accessed Acumen modules hosted on a HP DL 380 application server via Terminal Services during this performance trial
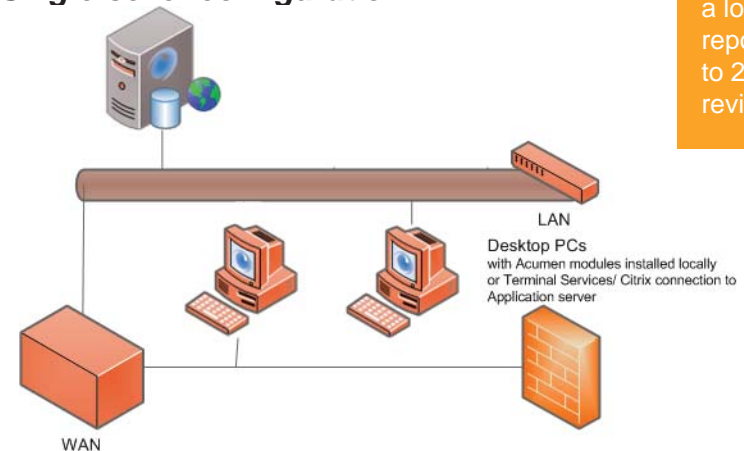
Server configuration options

Acumen leverages a three tier architecture consisting of web, database and application layers. These tiers may be deployed across multiple servers or a single server.

Multi-sever configuration

Application server
Database server SQLServer 2000
Web services server IIS 6.0
SAN (optional)
LAN
Desktop PCs with Acumen modules installed locally or Terminal Services/ Citrix connection to Application server
WAN

Single-sever configuration

LAN
Desktop PCs with Acumen modules installed locally or Terminal Services/ Citrix connection to Application server
WAN

Both multi-sever and single server configurations were tested at the HP PTAC, with single server configurations performing as well as a multi-server configuration using a local disk for file repository with up to 20 simultaneous reviewers.

Hardware specificiations

Acumen performance trials at HP's PTAC Center focused on a three-server and single-server configuration. This sheet summarizes hardware specs of the three server environment:

Application Server on HP DL 380
- CPU: 2 x Xeon 2.80GHz / 800MHz (Dual-Core, 2x2MB L2)
- Memory: 4GB Memory, PC2-3200 DDR2 SDRAM running at 400MHz with Advanced ECC and Online Spare capabilities. The memory subsystem on the DL380 G4 provides speeds up to 400MHz to transfer data at 6.4GB/s.
- Local disk: 4x36GB 15K RPM (two logical drives, each has 36GB with RAID-1)
- OS: Windows 2003 SP1, Server edition preferred
- MS Office Professional 2003
- IIS 6.0 with FrontPage extensions

Web Services Server on HP DL 380

- CPU: 2 x Xeon 2.80GHz / 800MHz (Dual-Core, 2x2MB L2)
- Memory: 4GB Memory, PC2-3200 DDR2 SDRAM running at 400MHz
- Local disk: 2x36GB 15K RPM (36GB logical drive with RAID-1)
- SAN: MSA 1000 w/512mb cache, 14 drives x 36gb 10k Disks, & Imbedded 2gb fc switch (two logical drives, each has 200GB with RAID-0). Alternately, EVA6000 2-controllers, 512 cache, 56 drives x 72gb @ 15k (max ~4TB) (two logical drives, each has 200GB with RAID-0)
- OS: Windows 2003 SP1, Server edition preferred
- MS Office Professional 2003
- IIS 6.0 with FrontPage extensions

Database Server on HP DL 380
- CPU: 2 x Xeon 2.80GHz / 800MHz (Dual-Core, 2x2MB L2)
- Memory: 4GB Memory, PC2-3200 DDR2 SDRAM running at 400MHz
- Local disk: 4x36GB 15K RPM (two logical drives, each has 36GB with RAID-1)
- OS: Windows 2003 SP1, Server edition preferred
- SQL Server 2000, Enterprise edition preferred

Acumen benefited from significant import and production performance improvements on HP's DL 380 mutli-sever environment by

- adding a second fiber channel path to the MSA 1000 SAN connected to the web services server

- increasing the amount of SAN cache from 256 MB to 512 MB

- creating an ethernet crossover connection between the database and web services/ file repository server.

Acumen Import - results

In Acumen's Import module, files are inventoried and documented with custodian information. All metadata and test is extracted and the MD5 hash value is assigned. Compound files are exploded, de-duplicated and optionally filtered by search terms.

Acumen imported up to 635 MB per hour on the multi-server environment. A 1 GB Outlook pst file comprised of 23,000+ files took 1 hour 42 minutes to import using the

**Acumen Import
in megabytes/ hour with various repository configurations**

Chart, Megabytes per hour:
- Import, single server environment: 623
- Import, local drive repository, multi-server environment: 574
- Import, SAN respository, multi-server environment: 502
- Import, crossover connection between database and webserver with SAN repository, multi-server environment: 635
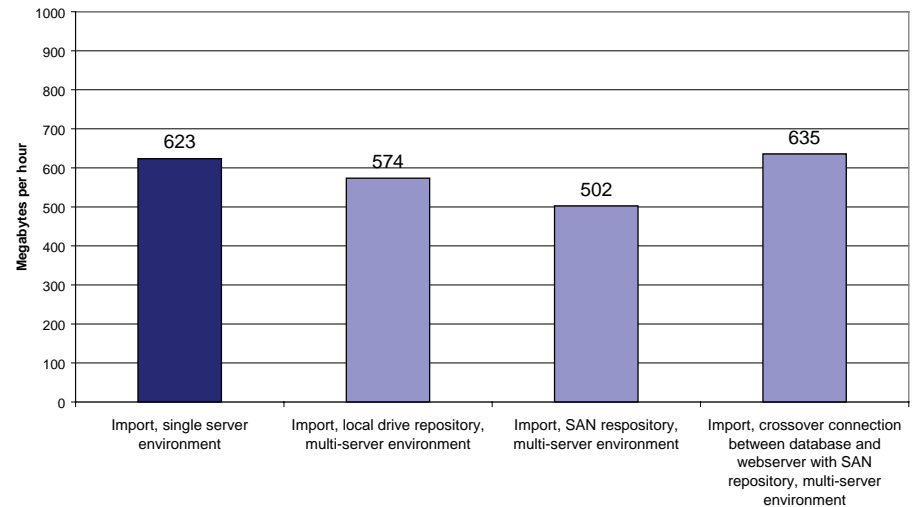
Repository variations
Three trials varied the location or configuration of the file repository. In the first trial, the file repository was located on the local (e:\) drive of the web services server and loaded 1 GB of pst data in 1 hr 53 min.   In the second trial, the file repository was located on an MSA 1000 SAN connected to the web services server and was about 9 percent slower than the local repository configuration. The third trial, however, showed a 26 percent performance improvement for Import with the file repository still located on SAN by creating an Ethernet connection crossover between the database server and the web services server connected to the SAN.

Data characteristics
Trial data consisted of a 1 GB pst containing 23,797 files, of which approximately half (540 MB) were messages, and the other half consisted of ppt (180 MB), docs (150 MB), pdf (86 MB), xls (30), and other file types.
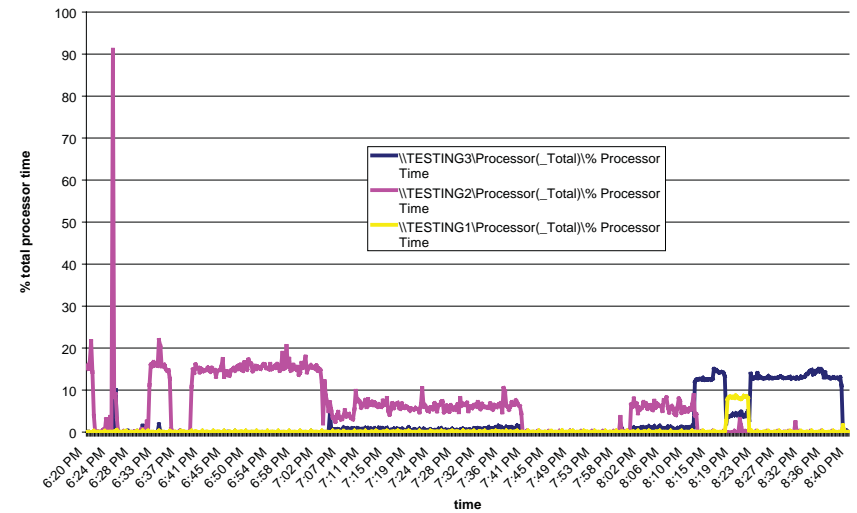
Acumen Import - single-server environment
CPU utilization with 1 GB pst
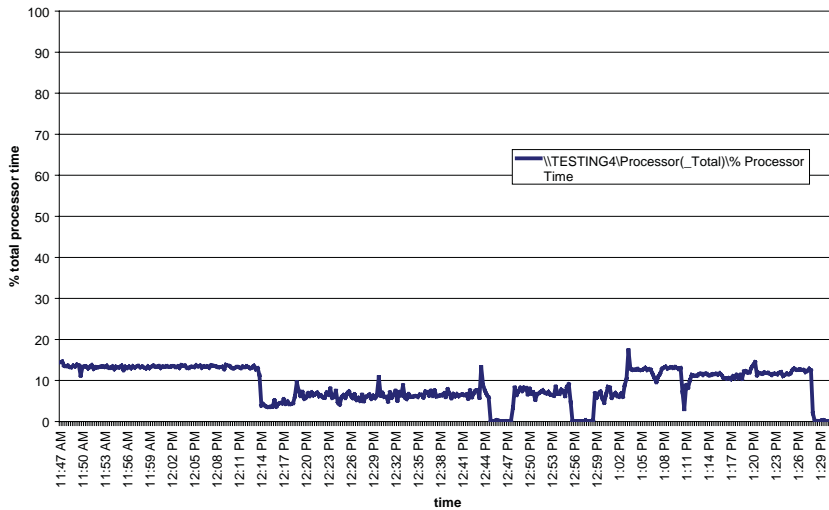
Capacity utilization on single-server environment

CPU utilization did not exceed 20% on the single server environment. Since import is a single-threaded process, due to limitations with Outlook profiles, subsequent tests published in this result set will explore creating additional virtual servers on the physical server in an effort to utilize more of the system's available processing capacity and memory.
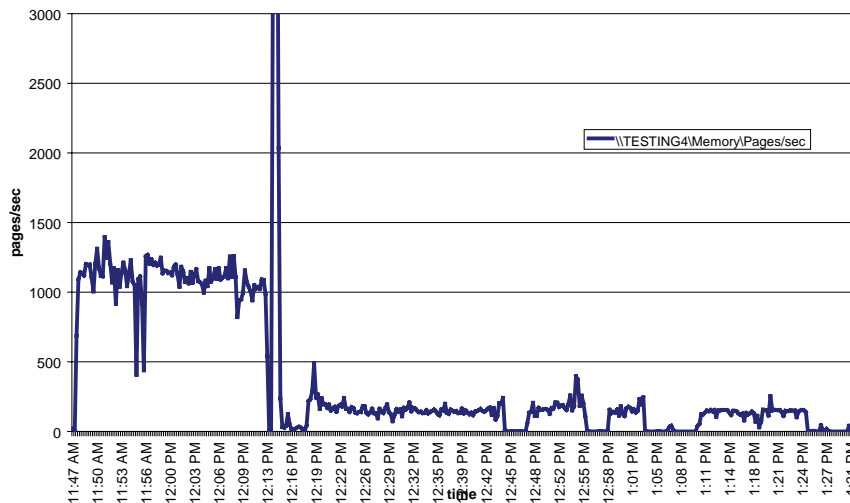
Acumen Import - single-server environment
Memory utilization with 1 GB pst

Acumen Import - multi-server environment
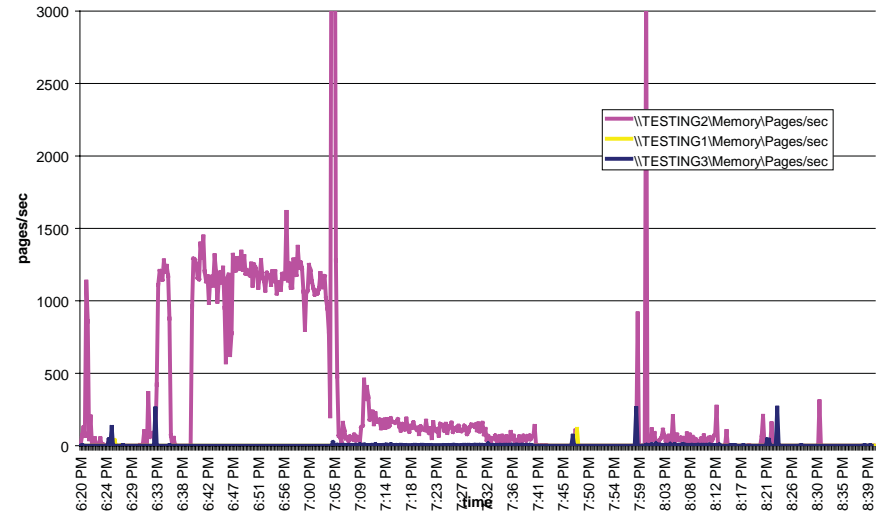## CPU utilization with 1 GB pst

Capacity utilization on multi-server environment

CPU utilization did not reach capacity on any one of the three servers. The application server (labeled as TESTING2 in the graph) approached and sustained approximately 20% CPU utilization during the local phases of import, while web server (labeled as TESTING3 in the graph) sustained 15% CPU utilization during the later de duplication and indexing phase of import.

Acumen Import - multi-server environment
Memory utilization with 1 GB pst

Acumen Production - results

Acumen produces responsive documents for either native file or image file export. The results that follow measure Acumen's performance with the production process running on server (the web services server); however, Acumen may be configured to run productions processes on multiple servers simultaneously.

Acumen tiff'ed up to 3737 pages per hour for a set comprised of mixed file types.

Data characteristics

Files were tiff'ed using Black and White Group 4 compression, although this setting may be changed in Acumen at the time of production to use LZW or other compressions. Producing files as PDF format rather than tif images performed 7% more slowly, all other things constant. Production performance varied based on the file types contained in the production set. See graph above for results.

Acumen Production (tiff and pdf)
Pages per hour with various repository configurations

Acumen Production (tiff)
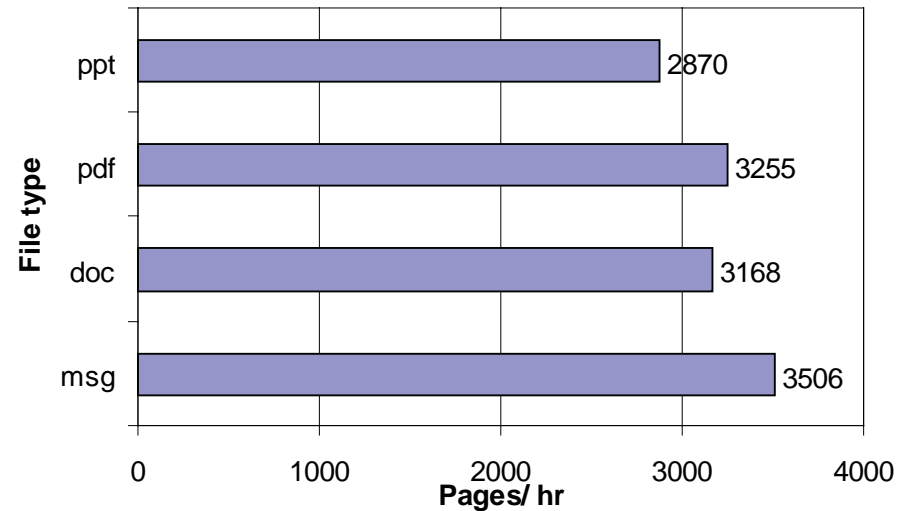**Pages per hour by native file type**
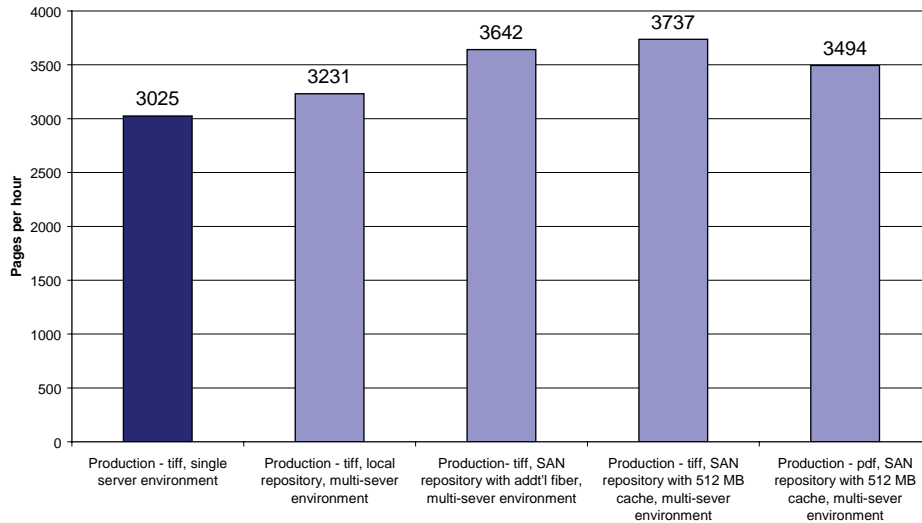
Repository variations

The single server environment was approximately 7% slower at producing tiff images as the closest comparable multi-server file repository configuration with local disk.

For the multi-sever environment, three trials varied the location and configuration of the file repository. In the first trial, a production set with 2K files (7162 pages) completed tiff'ing in 2 hrs 13 minutes while Acumen's file repository was hosted on a local drive. In the second trial, the same production set completed in 1 hr 58 minutes when Acumen's file repository was hosted on SAN with two fiber channels. Production performance was

Acumen Review - results

Acumen's site license includes access for 50 named reviewers. In the Review module user can search, save searches, print, redact, tag, annotate, designate, break or maintain parent-child relationships, QC and QA documents, and choose production formats.

To simulate a concurrent user load of 50 reviewers, testers logged into the review module with 50 unique users (reviewer1@ ceralon.com, reviewer2@ceralon.com... reviewer20@ceralon.com) at approximately the same time. Reviewers then worked through a series of test steps including designation, conversion, redaction in the drill-down view, and then marking the set as reviewed, QC'd, and QA'd to move it through the application workflow.

Acumen's Review module demonstrated the ability to support at least 50 named users authorized with a site license. All tested functions within the review module performed with 5 seconds or less response time with a load of up to 35 concurrent reviewers.

Acumen Review
CPU utilization with 50 concurrent users

Legend:
- \\TESTING3\Processor(_Total)\% Processor Time
- \\TESTING2\Processor(_Total)\% Processor Time
- Users
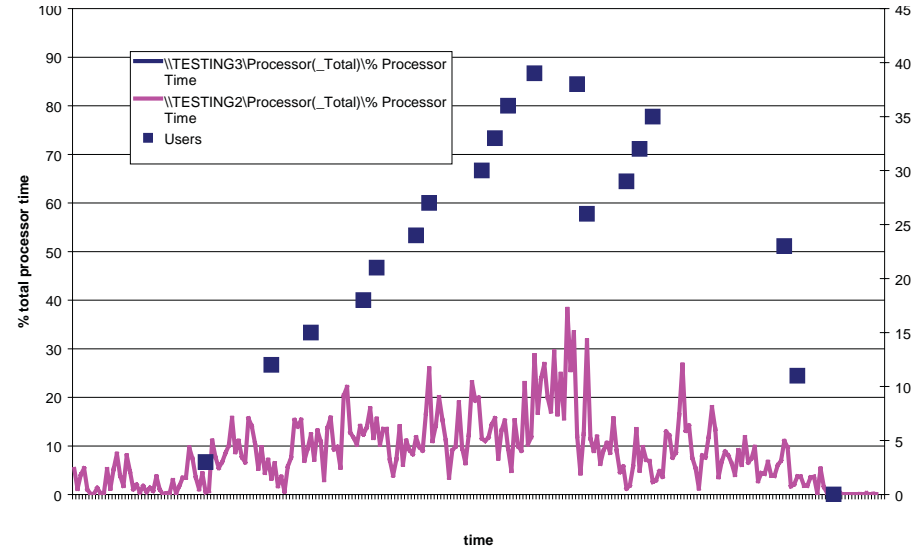
Acumen Review
Memory utilization with 50 concurrent users

Legend:
- \\TESTING2\Memory\Pages/sec
- \\TESTING1\Memory\Pages/sec
- \\TESTING3\Memory\Pages/sec
- Users