

Wide Area Information System in the Ariadne Network

Wide Area Information System στο Δίκτυο ΑΡΙΑΔΝΗ

A.S. DRIGAS
Applied Technologies Department
NCSR "DEMOKRITOS"
Ag. Paraskevi
GREECE
dr@imm.demokritos.gr
<http://imm.demokritos.gr>

1. Εισαγωγή

Η ευρεία εξάπλωση των δικτύων ηλεκτρονικών υπολογιστών σε συνδυασμό με τη δυνατότητα που παρέχεται από τις σύγχρονες εφαρμογές της πληροφορικής για την αποθήκευση και επεξεργασία πολύ μεγάλης ποσότητας πληροφορίας σε ψηφιακή μορφή, έχει μεταβάλει δραστικά τον τρόπο που ο άνθρωπος διανέμει, επεξεργάζεται, ανακαλεί και γενικά διαχειρίζεται την αποθηκευμένη πληροφορία. Η μεγάλη αύξηση της υπολογιστικής ισχύος και της χωρητικότητας των μονάδων αποθήκευσης σε συνάρτηση με τη θεαματική πτώση του κόστους ανά mip και mbyte αντίστοιχα την τελευταία τριετία οδήγησε σε αντίθεση με το παρελθόν στην "τοπική" αποθήκευση της πληροφορίας και τη διανομή, επεξεργασία, ακόμη και διαχείρισή της μέσω των δικτύων υπολογιστών σε, από κάθε ενδιαφερόμενο.

Η διαδικασία πρόσβασης, σε βάσεις πληροφοριών δηλαδή οι λειτουργίες αναζήτησης και μεταφοράς της πληροφορίας που βρίσκεται αποθηκευμένη σε ψηφιακή μορφή δεν είχε αλλάξει πολύ πριν από την τελευταία τριετία, από τον παραδοσιακό τρόπο πρόσβασης στις παραδοσιακές βιβλιοθήκες, κάνοντας χρονοβόρα τη διαδικασία αναζήτησης όταν η ποσότητα της αποθηκευμένης πληροφορίας γίνεται αρκετά μεγάλη. Στην περίπτωση αφ' ετέρου που η πρόσβαση στην πληροφορία γίνεται από μακριά μέσω των δικτύων υπολογιστών που παρεμβάλλονται μεταξύ του χρήστη και της βάσης πληροφοριών το πρωτόκολλο της πρόσβασης γίνεται πολύπλοκο και πρέπει να περιγράφει με σαφήνεια τα βήματα που θα ακολουθηθούν, για την απόκτηση της πληροφορίας όπως και όλες τις καταστάσεις που θα αντιμετωπιστούν κατά τη λειτουργία του. Η κατάσταση αυτή οδήγησε την προσπάθεια στην ανάπτυξη πρωτοκόλλων τα οποία υποστηρίζουν τη λειτουργία μεγάλων ηλεκτρονικών βάσεων πληροφοριών και παράλληλα το τρόπο πρόσβασης σε αυτές μέσω δικτύων. Ένα τέτοιο πρωτόκολλο είναι το Z39.50 [25] που σχεδιάστηκε από τον Dow Jones για τη διαχείριση ηλεκτρονικών βιβλιοθηκών.

Το πρωτόκολλο Z39.50 αποτελεί τον πυρήνα για ένα πακέτο προγραμμάτων που αναπτύχθηκε για να παρέχει ένα φιλικό περιβάλλον στο χρήστη που επιθυμεί την πρόσβαση σε ηλεκτρονικές βιβλιοθήκες μεγάλης χωρητικότητας μέσω των διαφόρων μέσων ψηφιακής επικοινωνίας όπως είναι τα IP, X25 και UUCP δίκτυα υπολογιστών. Ακόμη το πακέτο αυτό των προγραμμάτων παρέχει ένα εύχρηστο και με πολλές διαχειριστικές ευκολίες περιβάλλον για τη διαχείριση μεγάλων βάσεων πληροφοριών. Το πακέτο αυτό ονομάζεται WAIS (Wide Area Information System) [26], αναπτύχθηκε από την Thinking Machines Corporation και αποτελείται από μία σειρά προγραμμάτων που βοηθούν στην προσπάθεια για εύκολη πρόσβαση, γρήγορη αναζήτηση καθώς και σωστή διαχείριση μεγάλων ηλεκτρονικών βάσεων πληροφοριών.

Αναγνωρίζοντας τη λειτουργική αξία του προγράμματος και επιθυμώντας την αξιοποίηση του παγκόσμιου Δικτύου WAIS που εν το μεταξύ αναπτύχθηκε το δίκτυο Αριάδνη εγκατέστησε το πακέτο αυτό σε μία από τις μηχανές του προκειμένου να δώσει τη δυνατότητα στους χρήστες του να επικοινωνούν με τις βάσεις πληροφοριών που υποστηρίζονται από το πρόγραμμα αυτό στο παγκόσμιο Δίκτυο WAIS αλλά και με σκοπό στο μέλλον να δημιουργήσει μία τοπική βιβλιοθήκη που θα στηρίζεται στο εν λόγω πακέτο.

Στην παρούσα αναφορά περιγράφεται η δομή του πακέτου WAIS και η διαδικασία εγκατάστασης του σε μηχανή του δικτύου ΑΡΙΑΔΝΗ. Ακόμη γίνεται παρουσίαση για τον ειδικά σχεδιασμένο τρόπο με τον οποίο δομείται η πληροφορία μέσα στο WAIS ώστε η ευκολία και ταχύτητα αναζήτησης να μην επηρεάζεται δυσμενώς από το μέγεθος της όπως επίσης και από τον τρόπο που γίνεται η αναζήτηση και η ανάκτησή της.

2. Information Servers

Οι γενικές αρχές του μοντέλου λειτουργίας του πρωτοκόλλου αναζήτησης του WAIS περιγράφονται στις παραγράφους που ακολουθούν.

Μια τεχνική η οποία χρησιμοποιείται από μία ομάδα information servers και η οποία επιτρέπει την πρόσβαση σε μεγάλες συλλογές από αρχεία βασίζεται στα περιεχόμενα των αρχείων και τις ομοιότητες που εμφανίζουν αυτά μεταξύ τους. Η τεχνική αυτή στην οποία βασίζεται και ο information server WAIS ονομάζεται "Content Navigation". Χρησιμοποιώντας την τεχνική αυτή τα αρχεία μπορούν να ανευρεθούν ξεκινώντας από απλές ερωτήσεις στα Αγγλικά. Η αρχική ερώτηση μπορεί να είναι η επικεφαλίδα ή μια οποιαδήποτε λέξη κλειδί η οποία χαρακτηρίζει το αρχείο. Τα αρχεία που θα ανευρεθούν αρχικά μπορούν να χρησιμοποιηθούν στην συνέχεια από τον information server WAIS για εκτενέστερη αναζήτηση αρχείων όμοιων με αυτά.

Για κάθε ένα από τα αρχεία που ικανοποιούν την ερώτηση του χρήστη μπορεί να οριστεί ένα μέτρο για το κατά πόσο καλά ταιριάζει με την ερώτηση που έγινε (ποσοστό ικανοποίησης της ερώτησης ή αποκάλυψης του αρχείου). Το μέτρο αυτό είναι πολύ δύσκολο να εντοπιστεί και να λειτουργεί το ίδιο αποδοτικά σε διαφορετικούς τύπους αρχείων. Ωστόσο πρέπει να οριστεί ένας κατά το δυνατό εύκολος στη χρήση τρόπος με ευρεία εφαρμογή σε αρκετούς τύπους αρχείων που να αποδίδει το μέτρο

της αποκαλυπτότητας του αρχείου. Για τα αρχεία κειμένου ένα μέτρο θα μπορούσε να ήταν ο αριθμός των λέξεων που συμπίπτουν με τις λέξεις keywords των ερωτήσεων. Ο τρόπος αυτός δεν μπορεί να εφαρμοστεί σε αρχεία που δεν έχουν καθαρή μορφή κειμένου. Για άλλη μορφή πληροφορίας θα μπορούσε να χρησιμοποιηθεί άλλη μορφή μέτρου που να δίνει τη δυνατότητα αξιολόγησης της αποκαλυπτότητας του αρχείου. Έτσι η πληροφορία μπορεί να εντοπιστεί ρωτώντας τον "navigator" για αρχεία που περιέχουν μία ομάδα λέξεων. Τα αρχεία που το περιεχόμενο τους συμπίπτει περισσότερο με την ομάδα λέξεων ή με διαφορετικά λόγια η ομάδα λέξεων απαντάται περισσότερες φορές μέσα στο αρχείο, συγκεντρώνει μεγαλύτερη βαθμολογία και έτσι εμφανίζεται με σειρά προτεραιότητας στη λίστα των αρχείων που ικανοποιούν τα keywords.

Τα αρχεία που επιστρέφονται με τα μεγαλύτερα ποσοστά ικανοποίησης της ερώτησης μπορούν να χρησιμοποιηθούν για περαιτέρω επεξεργασία της αρχικής ερώτησης ώστε να προσδιοριστούν ακόμη περισσότερο αποκαλυπτικά αρχεία. Έτσι μπορεί να επιλεγεί ένα αρχείο που θεωρείται αρκετά αποκαλυπτικό και να ζητηθεί από τον server να βρει και άλλα αρχεία που να εμφανίζουν αρκετά κοινά σημεία με αυτό το συγκεκριμένο. Στην πραγματικότητα αυτό που χρησιμοποιεί ο server για να ανακαλύψει και άλλα αρχεία δεν είναι ένα κομμάτι αντίγραφο του θεωρούμενου ως αποκαλυπτικού αρχείου αλλά μία αναφορά από ομοιότητες μεταξύ των κειμένων ή διαφορετικά δείκτες που ορίζονται ως χαρακτηριστικοί για τα αρχεία. Η πληροφορία των ερωτήσεων προς το server και οι δείκτες των πιο αποκαλυπτικών αρχείων μπορούν να αποθηκευτούν τοπικά στο server ο οποίος θα συνεχίσει το ψάξιμο και αργότερα, κατά την εκτέλεση της διαδικασίας μεταφοράς προς τον client, από τον οποίο τέθηκαν τα ερωτήματα. Εδώ πρέπει να σημειωθεί ότι ολόκληρη η λειτουργία του WAIS στηρίζεται στο μοντέλο server client.

Υπάρχει επίσης και η δυνατότητα να δοθούν ορισμένα φίλτρα στον server κατά την επιλογή των κειμένων ώστε να μην μπαίνουν στην διαδικασία μεταφοράς τα αρχεία που επιπλέον περιέχουν πληροφορία που δε θεωρείται χρήσιμη.

Η διαδικασία μεταφοράς που χρησιμοποιείται από τους information servers είναι: Επιλέγεται ένα αρχείο ή μία λίστα από αρχεία που θεωρούνται ότι έχουν μεγάλο ποσοστό ικανοποίησης της ερώτησης και πρέπει να μεταφερθούν στον client για να αναγνωστούν και να επεξεργαστούν από το χρήστη. Τότε ο server μεταφέρει τα αρχεία στην μηχανή του client ώστε ο χρήστης πιο γρήγορα και πιο άνετα να χειριστεί την πληροφορία που βρίσκεται πλέον τοπικά. Για να αποφευχθεί άσκοπο γέμισμα του καναλιού από άχρηστη τελικά πληροφορία δηλαδή για να μην επιβαρυνθεί το δίκτυο αλλά και για να επιταχυνθεί η εργασία του χρήστη πρέπει να προσθέσει τα φίλτρα που προαναφέρθηκαν κατά την επεξεργασία και αναζήτηση στο server.

3. Δομή του WAIS

Το WAIS αναπτύχθηκε για να υποστηρίζει τη διανομή της πληροφορίας που βρίσκεται σε ηλεκτρονική μορφή μέσω των δικτύων

υπολογιστών. Το WAIS βελτιστοποιεί την εκμετάλλευση των δυνατοτήτων που δίνει η τεχνολογία υπολογιστών. Η πληροφορία μπορεί να είναι οτιδήποτε, από αρχεία κειμένου όπως επίσης ήχου ή και εικόνας. Από την πλευρά του χρήστη το πρόγραμμα αυτό σχεδιάστηκε για να είναι εύκολο στην χρήση του (φιλικό user interface) και με δυνατότητα σύνδεσης στο σύνολο των βάσεων πληροφοριών που στηρίζουν τη λειτουργία τους στο πρόγραμμα αυτό (παγκόσμιο Δίκτυο WAIS). Η επικοινωνία χρήστη προγράμματος γίνεται χρησιμοποιώντας σύμβολα της αγγλικής γλώσσας.

Το WAIS όπως και όλα σχεδόν τα προγράμματα που σχεδιάζονται για να λειτουργούν μέσω δικτύων υπολογιστών ακολουθούν το μοντέλο server client. Το πρόγραμμα αυτό ειδικότερα αποτελείται από τρία μέρη.

- a) server
- β) client
- γ) το πρωτόκολλο επικοινωνίας μεταξύ του server και του client.

Ο client περιέχει το user interface, το πρόγραμμα που αναλαμβάνει να δώσει στο χρήστη ένα φιλικό περιβάλλον ώστε ο χρήστης να δίνει τις εντολές του με τρόπο και γλώσσα κατανοητή από αυτόν και που επιπλέον λειτουργεί με menu. Ο client έπειτα αναλαμβάνει να μεταφράσει τις εντολές αυτές στο πρωτόκολλο επικοινωνίας για να τις στείλει στον server. Αυτός εκτελεί τη διαδικασία για την ανεύρεση της απάντησης (αρχεία με μεγάλο ποσοστό ικανοποίησης της ερώτησης) και στην συνέχεια την στέλνει πίσω στο χρήστη.

Τα στάδια που λαμβάνουν χώρα ώστε να συντελεστεί η προαναφερόμενη επικοινωνία είναι.

- 1) Ο χρήστης καλεί το πρόγραμμα του client.
- 2) Ο χρήστης θέτει τις ερωτήσεις για τον server χρησιμοποιώντας σύμβολα της αγγλικής γλώσσας.
- 3) Ο client μεταφράζει τις εντολές του χρήστη στο πρωτόκολλο επικοινωνίας και τις στέλνει στο server μέσω του δικτύου.
- 4) Ο server λαμβάνει τις εντολές του χρήστη επεξεργάζεται τα στοιχεία και επιστρέφει τα αποτελέσματα στον client μέσω του πρωτοκόλλου επικοινωνίας.
- 5) Ο client παρουσιάζει τα αποτελέσματα στο χρήστη και περιμένει νέα εντολή.

4. Client - User Interface

Το user interface είναι το βασικότερο πρόγραμμα του client. Ο χρήστης επικοινωνεί με το user interface και δε χρειάζεται να γνωρίζει τι ακριβώς συμβαίνει από εκεί και πέρα, δηλαδή πως το πρωτόκολλο διαχειρίζεται την εντολή του. Για την υποστήριξη των διαφόρων λειτουργικών συστημάτων έχουν σχεδιαστεί και αναπτυχθεί αρκετά user

interfaces για διάφορα λειτουργικά συστήματα και υπολογιστές όπως Unix, Vax, PC - MSDOS, Macintosh και για X-windows.

Αν και ο τρόπος με τον οποίο ο client επικοινωνεί με τον server είναι τελείως διάφανος από το χρήστη εντούτοις είναι ενδιαφέρον να παρουσιαστεί στο σημείο αυτό η δομή των ερωτήσεων που ο client στέλνει στο server.

Η δομή κάθε ερώτησης έχει τη μορφή.

```
(: question
  : version 2
  : seed words " XXXX"
  : relevant -documents()
  : sources
  ((: source-id
    : filename "wais-docs": )
   (: source-id
    : filename "CMNS" )
   )
  : result-documents ()
```

Περιγραφή των πεδίων:

Version : Ο αριθμός έκδοσης της δομής των ερωτήσεων που χρησιμοποιεί ο client του WAIS. Ο αριθμός αυτός είναι το 2.

seed words : Το string για το οποίο ψάχνεται η βάση πληροφοριών.

source <list> : Το πεδίο αυτό περιλαμβάνει αρχεία τα οποία ο χρήστης επιλέγει για την αναζήτηση των keywords.

relevant-documents () : Το πεδίο αυτό περιλαμβάνει τα αρχεία εκείνα που ο χρήστης θεωρεί αποκαλυπτικά για την πληροφορία που ψάχνει. Την πληροφορία αυτή αρκετοί servers τη χρησιμοποιούν για να βελτιώσουν την απόδοσή τους.

5. Server

Ο server δέχεται κλήσεις σε μία TCP πόρτα κατά σύμβαση στην πόρτα 210. Έχει τη δυνατότητα να τρέχει σαν standalone server αν κληθεί από το σύστημα την ώρα που σηκώνεται και η λειτουργία του μπει στο background και από τον inetd δαίμονα αν μπει κατάλληλη πρόταση στο αρχείο /etc/inet.conf .

Ο server έχει σχεδιαστεί για να παρέχει πολλές ευκολίες για τη διαχείριση της πληροφορίας. Έτσι εκτός του γεγονότος ότι ο server απαντά σε κλήσεις του client έχει και τη δυνατότητα να κρατά στατιστικά στοιχεία της χρήσης της βάσης πληροφοριών και να καταγράφει τα λάθη που συμβαίνουν κατά τη χρήση του προγράμματος. Μαζί με τον server δουλεύουν και προγράμματα που δημιουργούν την

ηλεκτρονική βάση πληροφοριών όπως το waisindex. Το wais μπορεί να χρησιμοποιηθεί και για εμπορικές βάσεις πληροφοριών. Για τον σκοπό μπορεί να χειρίζεται αρκετούς τρόπους χρέωσης.

Οι απαντήσεις που στέλνει ο server στον client είναι ουσιαστικά ένας αριθμός από αρχεία που περιέχονται στην βάση πληροφοριών και που εντοπίστηκαν στη διαδικασία αναζήτησης. Κάθε απάντηση του server στέλνεται υπό μορφή δομής στον client. Η δομή έχει τα ακόλουθα πεδία.

```
( : document-id
  : score      <integer>
  : start      <fragment>
  : end        <fragment>
  : document   <dogument>
)
```

```
( : fragment
  either
  : byte-pos   <integer>
  or
  : line-pos   <integer>
  or
  : para-id    <integer>
)
```

```
( : document
  : headline   <string>
  : doc-id     <doc-id>
  : source     <source-id>
  : number-of-lines <integer>
  : number-of-bytes <integer>
  : best-line   <integer>
  : type       <string>
  : date       <string>
)
```

```
( : doc-id
  : original-server      <any>
  : original-database    <any>
  : original-local-id    <any>
  : distributor-server   <any>
  : distributor-database <any>
  : distributor-local-id <any>
  : copyright-disposition <any>
)
```

6. Directory of servers

Στην ερώτηση πώς θα γνωρίζει ο client ποιοι άλλοι wais servers έχουν στηθεί και είναι προσβάσιμοι από τον κάθε χρήστη οι σχεδιαστές του wais έδωσαν απάντηση δημιουργώντας ένα κατάλογο όλων των servers που διατίθενται. Αυτός ο κατάλογος ανανεώνεται συνεχώς από τους κατασκευαστές του προγράμματος και είναι διαθέσιμος από πολλά FTP sites. Τον κατάλογο αυτόν τον ονόμασαν "directory of servers" γιατί είναι ακριβώς αυτό, ένα directory όπου κάθε αρχείο μέσα σε αυτό είναι μία εγγραφή για κάθε server.

Η πληροφορία σε κάθε αρχείο τοποθετείται σε δομή με την ακόλουθη μορφή.

```
( : source
  : version 3
  : ip-name "adonis.clnsnet.ariadne-t.gr"
  : ip-address "143.233.240.10"
  : ip-port 210
  : maintainer "postmaster@adonis.clnsnet.ariadne-t.gr"
  : database-wais : "directory of servers"
  : cost : 0.00
  : cost-unit : free
  : discription "xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx"
```

Περιγραφή των πεδίων :

version <integer> : Ο αριθμός της version είναι ο αριθμός εκδόσεως της δομής που ακολουθεί. Η δομή που περιγράφουμε έχει αριθμό version 3.

database-name <string> : Το string δηλώνει το όνομα της βάσης πληροφοριών στο server.

cost <float> : Το κόστος χρήσης του server. Στο πεδίο αυτό οι εμπορικές βάσεις πληροφοριών ανακοινώνουν το κόστος χρέωσης για τη χρήση του server. Η μονάδα μέτρησης ορίζεται στο επόμενο πεδίο.

cost-unit <cost-unit-type> : Στο πεδίο αυτό μπορούν να οριστούν τα παρακάτω.

```
: free
: dollars-per-session
: dollars-per-minutes
: dollars-per-retrieval
: dollars-per-query
: other
```

tcp-port <integer> : Στο πεδίο αυτό δηλώνεται ο αριθμός της πόρτας που χρησιμοποιείται για την επικοινωνία client server. Η πόρτα αυτή κατά σύμβαση είναι η 210.

7. Διαδικασία εγκατάστασης του WAIS

7.1. Γενικά

Η διαδικασία εγκατάστασης του WAIS αρχίζει με τη μεταφορά του πακέτου μέσω anonymous ftp από τη μηχανή με domain name think.com.

Το πακέτο αυτό βρίσκεται στο directory /wais με τη μορφή wais-8-b5.tar.Z. Στο directory αυτό υπάρχουν και άλλα αρχεία πληροφοριακά για το πρόγραμμα wais και προγράμματα για αρκετούς clients.

Η μεταφορά του αρχείου αυτού γίνεται σε Binary mode.

Μόλις γίνει η μεταφορά του προγράμματος σε περιοχή της μηχανής που θα εγκατασταθεί εκτελούμε τα ακόλουθα.

```
% uncompress wais-8-b5.tar.Z
```

```
% tar -xvf was-8-b5.tar
```

```
% cd wais-8-b5
```

Τα αρχεία που δημιουργούνται από τη διαδικασία αυτή είναι.

Makefile Το Makefile που κτίζει όλο το υπόλοιπο σύστημα.

README

INSTALLATION

RELEASE-NOTES

bin/ Το directory στο οποίο τοποθετούνται όλα τα εκτελέσιμα αρχεία.

doc/ Το directory στο οποίο βρίσκονται τα man pages και άλλα πληροφοριακά αρχεία.

ir / Το directory που περιέχει τον source κώδικα του waisserver και του waisindex.

ui/ Το directory που περιέχει τον source κώδικα για ορισμένα user interfaces.

wais-sources/ Το directory αυτό περιέχει ένα δείγμα βάσης πληροφοριών που δημιουργεί το wais κατά την εγκατάσταση του και τοποθετεί σε αυτήν όλα τα documentations.

x/ Το directory αυτό περιέχει τον source κώδικα ενός client για X11R4.

Έπειτα από την αποκατάσταση των αρχείων που γίνεται με τις εντολές uncompress και tar μπορεί να συνεχιστεί η διαδικασία εγκατάστασης.

Το πακέτο διανέμεται σε μορφή source κώδικα και πρέπει να μεταγλωτιστεί στη μηχανή που θα εγκατασταθεί, για να προσαρμοστεί στη αρχιτεκτονική και το λειτουργικό της συγκεκριμένης μηχανής.

Πριν ξεκινήσει η διαδικασία για τη δημιουργία των εκτελέσιμων προγραμμάτων πρέπει να διαβαστούν τα αρχεία README και INSTALLATION που δίνουν πληροφορίες για την εγκατάσταση του προγράμματος στους διάφορους τύπους μηχανών. Ακολουθώντας τις εντολές που δίνει ο κατασκευαστής του προγράμματος για την ιδιαιτερότητα της μηχανής που θα εγκατασταθεί γίνονται οι απαραίτητες αλλαγές στο Makefile. Το πρόγραμμα make του λειτουργικού συστήματος

αναλαμβάνει να μεταγλωτίσει και να ενώσει τον source κώδικα σύμφωνα με τη διαδικασία που περιγράφεται στο Makefile.

Όταν το make τελειώσει επιτυχώς τα εκτελέσιμα αρχεία μεταφέρονται στις μόνιμες θέσεις τους. Τα χρήσιμα αρχεία είναι.

```
waisserver waisserver.d
waisq
waissearch
waisindex
swais
```

Από τα αρχεία αυτά τα **waisserver** και **waisserver.d** που είναι οι servers εγκαθίστανται στο directory **/etc**

Τα αρχεία **waisindex**, **waisq**, **waissearch** και **swais** εγκαθίστανται στο directory **/usr/local/bin** για να βρίσκονται αυτόματα στο path κάθε χρήστη.

Τα **man pages** τοποθετούνται στο directory **/usr/local/man**.

7.2. Εγκατάσταση του Server

Ο server- daemon μπορεί να καλείται από το σύστημα την ώρα που σηκώνεται και να τρέχει καθ' όλη τη διάρκεια που το σύστημα είναι σε λειτουργία.

Για το λόγο αυτό βάζουμε την ακόλουθη πρόταση στο rc.local file.

```
#
# WAIS SERVER
#
echo -n "wais server :">/dev/console
if [ -f /etc/waisserver ]; then
    /etc/waisserver -p 210 -d /home/wais/wais-source \
    > /home/wais/server.log 2>&1 &
    echo -n 'done'
#
```

Επίσης ο server έχει τη δυνατότητα να τρέχει όταν κληθεί στην TCP/IP πόρτα που δέχεται τις κλήσεις, όταν έχουν τοποθετηθεί οι ακόλουθες προτάσεις στα αρχεία.

Στο αρχείο **/etc/services**

```
z3950          210/tcp          # z3950 protocol for wais
```

Στο αρχείο **/etc/inetd.conf**

```
z3950 stream tcp nowait root swais /etc/waisserver.d
```

Ο waisserver δέχεται τις ακόλουθες επιλογές.

```
p [n] Η TCP στην οποία δέχεται τις κλήσεις ο server.
```

d [directory] Το directory που βρίσκονται τα αρχεία της πληροφορίας.

e [filename] Όπου filename το όνομα του αρχείου που αποθηκεύονται τα errors που συμβαίνουν κατά την κλήση του wais.

l [log_level] Η επιλογή αυτή θέτει το επίπεδο των λαθών που θα καταχωρούνται στο αρχείο που η επιλογή -e προσδιορίζει. Τα log_levels που είναι διαθέσιμα είναι τα 0, 1, 5, 10. Το log_level ίσον 0 δεν κρατά τα λάθη που συμβαίνουν. Το log_level ίσο με 1 κρατά μόνο τα λάθη με την υψηλότερη προτεραιότητα. Το log_level ίσον 5 κρατά και τα λάθη με μέση προτεραιότητα. Το log_level ίσον με 10 κρατά όλα τα λάθη που συμβαίνουν.

u user Η επιλογή αυτή αναγκάζει τον server να θέτει ένα συγκεκριμένο user id όταν ανοίγει την κλήση.

v Τυπώνει πληροφορίες για τον server.

7.3. Δημιουργία της βάσης πληροφοριών

Το wais χρησιμοποιεί ένα δικό του τρόπο για να βρίσκει τα κείμενα που περιέχουν τις λέξεις ή τις προτάσεις για τις οποίες ο χρήστης ψάχνει μέσω αυτού. Η ειδική κωδικοποίηση που δημιουργεί το wais δεσμεύει χώρο δίσκου ίσο με το μέγεθος των αρχείων της βάσης πληροφοριών. Η κωδικοποίηση αυτή γίνεται με τη χρήση της εντολής waisindex. Η σύνταξη της εντολής waisindex είναι.

waisindex -d [index-filename] [-a] filename filename ...

Οι παράμετροι που δέχεται η εντολή αυτή είναι:

-d [file-name] Το όνομα της βάσης που θα δημιουργηθεί

-a Η παράμετρος αυτή χρησιμοποιείται όταν χρειάζεται να γίνει προσθήκη (append) στην ήδη υπάρχουσα βάση πληροφοριών.

-mem Με την παράμετρο αυτή προσδιορίζεται η ποσότητα της μνήμης του συστήματος που θα χρησιμοποιηθεί για το indexing που δημιουργεί η εντολή. Η επιλογή αυτή έχει χρησιμότητα όταν πρόκειται να επεξεργαστεί μεγάλη ποσότητα πληροφορίας.

-register Η επιλογή αυτή ανακοινώνει στο directory of servers την ύπαρξη της βάσης πληροφοριών που μόλις δημιουργήθηκε.

-export Η επιλογή αυτή ανακοινώνει στο directory of servers τη διεύθυνση του υπολογιστικού συστήματος που την υποστηρίζει και την TCP πόρτα στην οποία δέχεται κλήσεις από το client. Η επιλογή -register δεν προσδιορίζει τα στοιχεία της διεύθυνσης και της TCP πόρτας. Με τη χρήση των -register και -export κατά την κλήση του waisindex αυτόματα στέλνεται mail στη διεύθυνση wais-directory-of-servers@quake.think.com

για να καταχωρηθεί η βάση πληροφοριών στο directory of servers. Για το λόγο αυτό η think.com αν και ενθαρρύνει τις ανακοινώσεις στο directory of servers προτρέπει αυτό να γίνεται όταν υπάρχει ουσιαστική πληροφορία για αυτό το σκοπό.

Η εντολή waisindex επεξεργάζεται τα αρχεία της πληροφορίας και δημιουργεί τα ακόλουθα βοηθητικά αρχεία για να γίνεται πιο γρήγορα η αναζήτηση στη βάση.

Τα αρχεία αυτά είναι:

[database-name].cat : Το αρχείο αυτό περιέχει ένα κατάλογο όλων των αρχείων που επεξεργάστηκαν από το waisindex για να προκύψει η βάση πληροφοριών.

[database-name].dct : Το αρχείο αυτό είναι ένα index file .

[database-name].doc

[database-name].hl

[database-name].inv

8. User Interface

Το wais user interface για το λειτουργικό σύστημα UNIX είναι το εκτελέσιμο αρχείο swais. Το πρόγραμμα αυτό κρίνεται σκόπιμο να τοποθετείται πάντα σε directory που βρίσκεται κατά σύμβαση στο path κάθε χρήστη.

Στο interface παρουσιάζεται ένας κατάλογος όλων των μηχανών που υποστηρίζουν τον wais server και δίπλα στο όνομα κάθε wais server δίνονται συνοπτικές πληροφορίες για το είδος της πληροφορίας που παρέχεται από τη βάση, καθώς και το κόστος για τη χρήση της βάσης αν υπάρχει τέτοιο.

Wais #	Server	Source Selection Source	Sources: 445
001:	[archie.au]	aarnet-resource-guide	Free
002:	[munin.ub2.lu.se]	academic_email_conf	Free
003:	[wraith.cs.uow.edu.au]	acronyms	Free
004:	[archive.orst.edu]	aeronautics	Free
005:	[ftp.cs.colorado.edu]	aftp-colorado-adu	Free
006:	[nostromo.oes.orst.edu]	agricultural-market-news	Free
007:	[archive.orst.edu]	alt.drugs	Free
008:	[wais.oit.unc.edu]	alt.gopher	Free
009:	[sun-wais.oit.unc.edu]	alt.sys.sun	Free
010:	[wais.oit.unc.edu]	alt.wais	Free
011:	[alfred.ccs.carleton.]	amiga-slip	Free
012:	[munin.ub2.lu.se]	amiga_fish_contents	Free
013:	[coombs.anu.edu.au]	ANU-Aboriginal-Studies	\$00.00/min
014:	[coombs.anu.edu.au]	ANU-Asian-Computing	\$00.00/min
015:	[coombs.anu.edu.au]	ANU-Asian-Religions	\$00.00/min
016:	[coombs.anu.edu.au]	ANU-CAUT-Projects	\$00.00/min
017:	[coombs.anu.edu.au]	ANU-Coombspapers-Index	\$00.00/min
018:	[coombs.anu.edu.au]	ANU-Local-Waservers	\$00.00/min
019:	[coombs.anu.edu.au]	ANU-SSDA-Catalogues	\$00.00/min
020:	[coombs.anu.edu.au]	ANU-Thai-Yunnan	\$00.00/min
021:	[coombs.anu.edu.au]	ANU-Theses-Abstracts	\$00.00/min
022:	[quake.think.com]	Applications-Navigator	Free

023:	[info.cin.salford.ac.]	Applied-Science-and-Technol.	Free
024:	[info.cin.salford.ac.]	Applied-Science-and-Technol.	Free
025:	[132.183.190.21]	Arabiopsis-BioSci	Free
026:	[weeds.mgh.harvard.edu]	Arabiopsis_thaliana_Genome	Free
027:	[ftp.tex.ac.uk]	archeological_computing	Free
028:	[archive.orst.edu]	archie-orst.edu	Free

keyword :

<space> selects, w for keywords, arrows move, <return> searches, q quits, or ?

Οι εντολές του swais client είναι.

j, down arrow ^N Move down one source
k, up arrow, ^P Move up one source
J, ^V, ^D Move down one screen
K, <esc> v, ^U Move up one screen
Position to source number ###
/sss Search for source sss
<space>, <period> Selection current source
= Deselection all sources
v, <comma> View current source info
<ret> Perform search
s Select new sources
w Select new keywords
X, - Remove current source permantly
o Set and show swais options
h, ? Show help
H Display program history
q Exit the program

9. Shell user interfaces

Τα εκτελέσιμα αρχεία waisq και waisearch είναι απλά user interfaces που μπορούν να χρησιμοποιηθούν είτε από το διαχειριστή της βάσης για έλεγχο της λειτουργίας του server είτε από οποιονδήποτε θέλει να αναζητήσει πληροφορία από συγκεκριμένη βάση και θέλει ένα γρήγορο και αρκετά δυνατό εργαλείο.

Το user interface waisq παίρνει τις παραμέτρους:

waisq [-h host-machine] [-p [service-or-port]] [-d database] [-m max-results] word word ...

Το user interfase waisq είναι λίγο πιο πολύπλοκο στην λειτουργία από το waissearch αλλά είναι πιο δυνατό εργαλείο. Το waisq δέχεται τις εξής παραμέτρους.

waisq
[-f question_file]
[-s sourcedir]
[-S sourcename]

```
[ -c common_sourcedir ]  
[ -m max_Result_Docs ]  
[ -v document_number ]  
[ -g ]  
[ -h ]  
[ -t ]  
[ -V ]  
[ word [ word [ word [ ... ] ] ] ]
```

όπου:

[-f question-file] : είναι το αρχείο στο οποίο θα αποθηκευτούν τα αποτελέσματα της ερώτησης προς τον waisserver. Αν δεν προσδιοριστεί το αρχείο αυτό τότε τα αποτελέσματα εμφανίζονται στην οθόνη του τερματικού.

[-s sourcedir] : είναι το directory στο οποίο βρίσκεται η βάση πληροφοριών. Αν το directory αυτό δεν προσδιοριστεί τότε θεωρείται ως default directory το ~/wais-sources/.

[-S sourcename] : είναι το όνομα της βάσης πληροφοριών.

[-m maxdocs] : προσδιορίζει το μέγιστο αριθμό των απαντήσεων που επιστρέφονται από μια ερώτηση.

[-v document number] : ο αριθμός του κειμένου που προσδιορίζεται από μία έγκυρη απάντηση και που επιλέγεται να αποκτηθεί κατά τη διαδικασία της ερώτησης.

[-V] : επιστρέφει τον αριθμό εκδόσεως του προγράμματος.