SqlOnWindows

You can use SQL to store your bayes and awl databases if you're running on Windows. Storing the database in SQL reportedly reduces cpu overhead and increases speed. You can likely use any database that has a DBD module and permits threads to share handles. ADO does not permit handle-sharing, so you may not use it. Others have successfully installed MySQL on Windows and used that. This how-to provides information on using Microsoft's data engine. It was tested with MSDE 2000.

1. Install the MSDE database engine. You can get it on the SQL 2000 service pack 4 download page: http://www.microsoft.com/downloads/details.aspx?familyid=8E2DFC8D-C20E-4446-99A9-B7F0213F8BC5&displaylang=en. You want the SQL2000.MSDE-KB884525-SP4-x86-ENU.EXE file near the bottom of the page. Extract it and follow the installation instructions. 2. Create a new database. Create a new login. Give the login rights to the database. I use the tools that came with SQL Server 2000 to do this. I don't know that Microsoft still has the tools available for MSDE 2000 since they want you to use SQL Express 2005 now. I haven't tested SQL Express 2005 mainly because our environment uses SQL Server 2000. 3. Create the RPAD function

```
if exists (select * from dbo.sysobjects where id = object_id(N'[dbo].[RPAD]') and xtype in (N'FN',
N'IF', N'TF'))
drop function [dbo].[RPAD]
GO

SET QUOTED_IDENTIFIER ON
GO
SET ANSI_NULLS ON
GO

CREATE function RPAD (@cString nvarchar(4000), @nLen smallint, @cPadCharacter nvarchar(4000) = ' ' )
returns nvarchar(4000)
as
   begin
       declare @length smallint, @lengthPadCharacter smallint
       if @cPadCharacter is NULL or  datalength(@cPadCharacter) = 0
          set @cPadCharacter = space(1)
       select  @length  = datalength(@cString)/(case SQL_VARIANT_PROPERTY(@cString,'BaseType') when
'nvarchar' then 2  else 1 end) -- for unicode
       select  @lengthPadCharacter  = datalength(@cPadCharacter)/(case SQL_VARIANT_PROPERTY
(@cPadCharacter,'BaseType') when 'nvarchar' then 2  else 1 end) -- for unicode

       if @length >= @nLen
          set  @cString = left(@cString, @nLen)
       else
          begin
             declare  @nRightLen smallint
             set @nRightLen  =  @nLen - @length -- Quantity of characters, added on the right
             set @cString =  @cString + left(replicate(@cPadCharacter, ceiling(@nRightLen
/@lengthPadCharacter) + 2), @nRightLen)
          end

     return (@cString)
    end

GO
SET QUOTED_IDENTIFIER OFF
GO
SET ANSI_NULLS ON
GO
```

3. Edit c:/perl/site/lib/mail/spamassassin/bayesstore/SQL.pm. Find the line that says

```
  return "RPAD(token, 5, ' ')";
```

and change it to

```
  return "dbo.RPAD(token, 5, ' ')";
```

3. Create the tables in the database.

```
CREATE TABLE [dbo].[awl] (
        [username] [varchar] (100) COLLATE SQL_Latin1_General_CP1_CI_AS NOT NULL ,
```

```
        [email] [varchar] (200) COLLATE SQL_Latin1_General_CP1_CI_AS NOT NULL ,
        [ip] [varchar] (10) COLLATE SQL_Latin1_General_CP1_CI_AS NOT NULL ,
        [count] [int] NOT NULL ,
        [totscore] [float] NOT NULL
) ON [PRIMARY];

CREATE UNIQUE CLUSTERED
  INDEX [PK_awl] ON [dbo].[awl] ([username], [email], [ip])
WITH
    DROP_EXISTING
ON [PRIMARY];


CREATE TABLE [dbo].[bayes_vars] (
        [id] [int] IDENTITY (1, 1) NOT NULL ,
        [username] [varchar] (200) COLLATE SQL_Latin1_General_CP1_CI_AS NOT NULL ,
        [spam_count] [int] NOT NULL ,
        [ham_count] [int] NOT NULL ,
        [token_count] [int] NOT NULL ,
        [last_expire] [int] NOT NULL ,
        [last_atime_delta] [int] NOT NULL ,
        [last_expire_reduce] [int] NOT NULL ,
        [oldest_token_age] [int] NOT NULL ,
        [newest_token_age] [int] NOT NULL
) ON [PRIMARY];

CREATE UNIQUE CLUSTERED
  INDEX [PK_bayes_vars] ON [dbo].[bayes_vars] ([id])
WITH
    DROP_EXISTING
ON [PRIMARY];

CREATE UNIQUE
  INDEX [bayes_vars_idx1] ON [dbo].[bayes_vars] ([username])
WITH
    IGNORE_DUP_KEY
    ,DROP_EXISTING
ON [PRIMARY];


CREATE TABLE [dbo].[bayes_expire] (
        [id] [int] NOT NULL ,
        [runtime] [int] NOT NULL
) ON [PRIMARY];

CREATE UNIQUE CLUSTERED
  INDEX [PK_bayes_expire] ON [dbo].[bayes_expire] ([id])
WITH
    DROP_EXISTING
ON [PRIMARY];


CREATE TABLE [dbo].[bayes_global_vars] (
        [variable] [varchar] (30) COLLATE SQL_Latin1_General_CP1_CI_AS NOT NULL ,
        [value] [varchar] (200) COLLATE SQL_Latin1_General_CP1_CI_AS NOT NULL
) ON [PRIMARY];

CREATE UNIQUE CLUSTERED
  INDEX [PK_bayes_global_vars] ON [dbo].[bayes_global_vars] ([variable])
WITH
    DROP_EXISTING
ON [PRIMARY];

INSERT INTO bayes_global_vars VALUES ('VERSION','3');


CREATE TABLE [dbo].[bayes_seen] (
        [id] [int] NOT NULL ,
        [msgid] [varbinary] (200) NOT NULL ,
        [flag] [char] (1) COLLATE SQL_Latin1_General_CP1_CI_AS NOT NULL
) ON [PRIMARY];
```

```
CREATE UNIQUE CLUSTERED
  INDEX [PK_bayes_seen] ON [dbo].[bayes_seen] ([id], [msgid])
WITH
    DROP_EXISTING
ON [PRIMARY];


CREATE TABLE [dbo].[bayes_token] (
        [id] [int] NOT NULL ,
        [token] [char] (5) COLLATE SQL_Latin1_General_CP1_CI_AS NOT NULL ,
        [spam_count] [int] NOT NULL ,
        [ham_count] [int] NOT NULL ,
        [atime] [int] NOT NULL
) ON [PRIMARY];

CREATE UNIQUE CLUSTERED
  INDEX [PK_bayes_token] ON [dbo].[bayes_token] ([id], [token])
WITH
    DROP_EXISTING
ON [PRIMARY];

CREATE
  INDEX [bayes_token_idx1] ON [dbo].[bayes_token] ([token])
WITH
    DROP_EXISTING
ON [PRIMARY];

CREATE
  INDEX [bayes_token_idx2] ON [dbo].[bayes_token] ([id], [atime])
WITH
    DROP_EXISTING
ON [PRIMARY];
```

4. Install the DBI and ODBC database modules.

```
ppm install DBI
ppm install DBD-ODBC
```

4. Edit local.cf to change the settings for AWL

```
# Use SQL
auto_whitelist_factory Mail::SpamAssassin::SQLBasedAddrList
# Use ODBC connector
user_awl_dsn DBI:ODBC:Driver={SQL Server};Server=localhost;Database=database-name
user_awl_sql_username database-user
user_awl_sql_password database-user-password
# Use this for global AWL
user_awl_sql_override_username global
```

5. Edit local.cf to change the settings for bayes

```
# Use SQL
bayes_store_module          Mail::SpamAssassin::BayesStore::SQL
# Use ODBC connector
bayes_sql_dsn               DBI:ODBC:Driver={SQL Server};Server=MAIL;Database=MailData
bayes_sql_username          MailServer
bayes_sql_password          mail*data
# Use this for global bayes
bayes_sql_override_username global
```