July 20, 2010

To: MOVES Task Force

From: Mohamed A. Khan

Mobile Sources Control Program
MD Department of the Environment

Subject: **MDE's experience with MOVES Model – Distributed Processing**

We all know that MOVES model run time can be long, especially for annual emissions or emission factors. EPA recommended distributed processing feature can be very useful in reducing the actual run times, deploying more than one computer. The concept is simple and here are the highlights of MOVES Distributed Processing.

● MOVES consists of two applications (By Default): "Master" and "Worker"

● Master and Worker communicate through text files in the "SharedWork" folder

● Work is bundled in "packets" and placed in the "SharedWork" folder that can be
   processed by any worker application

● Location = C:\ Program Files\MOVES20091221\SharedWork

● Two MOVES configuration files exist in the MOVES program folder:
   MOVESConfiguration.txt   &   WorkerConfiguration.txt

   Both Master and Worker configuration files have the same path, specifying the
   Master and the Worker applications where to put files, as follows:

   sharedDistributedFolderPath= C:\Program Files\ MOVES20091221\SharedWork

● Distributed Processing needs a folder on network, accessible by all computers

● MDE used the following path for distributed processing, ensuring that the
   new "SharedWork" folder has full read/write access and <u>all the computers share the
   same distributed folder path:</u>

   sharedDistributedFolderPath = L:\MobileSrcs\MS_Exchange\SharedWork

● Master need to be the fastest PC, with most memory and hard drive space

● Start a worker on each of the other computers included in the processing using the
   'Worker' desktop icon. Notice a Worker ID file for each computer which is running a
   Worker, in the shared distributed folder

● Click the 'Master' desktop icon to start the MOVES application. Once the Master is started, a Master ID file will appear in the shared distributed folder

● Run the MOVES application on the Master machine as normal. When the bundles are produced, they will appear in the shared distributed folder on the network. The Worker applications on the other machines will see the bundles, read them and process them. The results of their processing will be placed back in the shared distributed folder on the network to be picked up by the Master application and stored in the output database on the Master machine.

● Watch the creation of "to do" files, "in process" files and "done" files while MOVES is running. When MOVES is complete, the only remaining files in the shared distributed folder on the network will be the Worker and Master ID files.

● Worker computers may be disconnected and the system will automatically detect this and safely reroute data to the remaining workers, including data that was already being processed by a disconnected worker. However, if all workers are disconnected, no work will be done and the MOVES run will fail.

● Machines running Workers will use most of the available CPU for processing and it will be difficult to run other applications while processing MOVES bundles from the Master computer.

● **What is the gain in run-time?** It depends on the CPU speed and memory resources of the machines involved. Too many/too slow workers tend to show very minimal gains.

● Example: Baltimore Area 2007 Annual Scenario (VOC, CO, NOx) Run time

1) PC 'A'	2.0 GHz, 2.0 GB RAM ……………..	7.43 hours
2) PC 'B'	3.25/2.99 GHz, 3.25 GB RAM) …….	3.90 hours
	MOVES Run Time Gain ……	3.53 hours gain = 48%

3) Distributed Processing with PC 'B' & PC 'A'   **2.60 hours**

                              MOVES Run Time Gain ……   1.30 hours gain = 33%

4) Dist. Processing with PC 'B' & 2 units of PC 'A'   2.50 hours

                              MOVES Run Time Gain ……   1.40 hours gain = 36%

                    (Only 3% gain with one more PC added)

Better results possible if the Workers are also dual processor units with more RAM.