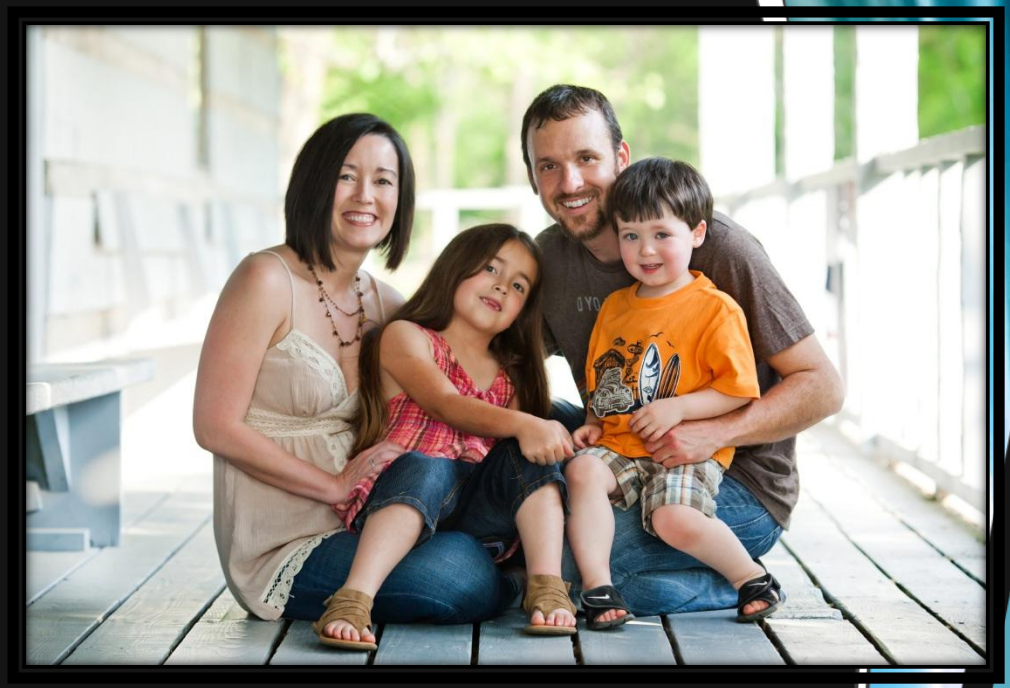
Performance Tuning in SSIS 2012

Hope Foley

Perpetual Technologies, Inc.

Who Am I?

- Microsoft Team Lead
- SQL Server DBA
- (MCITP: Database Administration in 2005 and 2008)
- In IT industry for 12 years. DBA for 6 years. I've been with PTI for 4 years.
- Worked for various industries such as large insurance companies, government entities, large and small corporations, hospitals and medical related businesses and on and on



Agenda

1. Intro
2. Go over monstrosity of a package
3. Baseline info
4. Concepts to make package perform better
5. Show much better version package
6. Additional list of considerations for performance
7. Introduce SSIS catalog
8. Go over new features of SSIS
9. Resources
10. Wrap it up

No magic button

SQL Server Integration Services 2012 is super awesome but sadly no magic button.

So will still have to write packages sensibly if you want them to perform well.

Look away…look away

Best way to learn what not to do is to try it for yourself.

Let’s take a look at a hideous monstrosity of an SSIS package.

Demo time!

Execution Differences

Prior to deploy (time)

Old Mess	Hip New	Time Saved	% Improvement
37:03.0	6:41.5	30:21.5	122%
38:41.1	2:29.1	36:12.0	107%
32:21.1	2:25.9	29:55.2	108%

After Deployed (seconds)

Old Mess	Hip New	Time Saved	% Improvement
2177.845	264.995	1912.85	88%
2067.65	142.792	1924.88	93%
2892.345	229.311	2663.034	92%

Numbers on an old Lenovo ThinkPad with 4GB RAM.

More Performance Tweaks to Consider

- Keep source info to minimum – don't need it…get it outta there
- Designed for parallelism
- Tasks to run parallel
- Avoid row by row tasks (SCD or OLE DB)
- Flat files use fast parse when can – saying my data is good in there and know it
- DefaultBufferMaxRows & DefaultBufferMaxSize
- BLOBTempStoragePath & BufferTempStoragePath
- BDD (Balanced Data Distributor – SQLCAT to release version 2012 near future)

SSIS Catalog

- What is it?
- What can catalog do for me?
- Logging levels – server level
- Data Taps
- And PowerShell just for fun

Demo time

Logging Levels

There are 4 logging levels you can set for the SSIS Catalog

Logging Level	Events Logged	Notes
None	None	Success/Fail and that's it
Basic	OnPreValidate OnPostValidate OnPreExecute OnPostExecute OnInformation OnWarning OnError	Similar to what seen when run in DTExec
Performance	OnWarning OnError	Required to get runtime information but doesn't log all events in basic
Verbose	Everything	Captures tons – careful as pretty safe to say will increase runtime

Data Taps

Allows you to “tap” into the data via T-SQL and extract data from any data flow. Will create a *.csv file for you to use. Can use to assist in data issues with development.

add_data_tap [@execution_id =] execution_id [@task_package_path =] task_package_path [@dataflow_path_id_string =] dataflow_path_id_string [@data_filename =] data_filename [@num_rows =] num_rows

exec catalog.add_data_tap @execution_id, '\Package\Insert Header', 'Paths[TaskName.OLE DB Source Output]', ‘datatapfilename.csv 5000

PowerShell

An Integration Services assembly is available with 2012. This will allow you to script out pretty much any of your SSIS functions into PowerShell scripts for automation.

Examples:

```
# Load the IntegrationServices Assembly $loadStatus =
[Reflection.Assembly]::Load("Microsoft"+
".SqlServer.Management.IntegrationServices" + ",
Version=11.0.0.0, Culture=neutral" + ",
PublicKeyToken=89845dcd8080cc91")

# Store the IntegrationServices Assembly namespace to avoid
typing it every time $ISNamespace =
"Microsoft.SqlServer.Management.IntegrationServices" Write-Host
"Connecting to server ..." # Create a connection to the server
$constr = "Data Source=localhost;Initial Catalog=master;Integrated
Security=SSPI;"
```

Honorable New Feature Mentions

Too Many to list but some of my favorites:

1. Parameters (required at runtime variables)
2. BIDSHelper (soon to be 2012 friendly from word on street)
3. Store multiple versions of packages
4. Undo/Redo
5. Expression decoration
6. CDC (For Oracle/Pull info/CDC splitter)
7. DQS (run data against knowledge base)

Resources

Matt Masson – SSIS development team.

http://blogs.msdn.com/b/mattm/

http://www.mattmasson.com/

Specific to troubleshooting

http://blogs.msdn.com/b/mattm/archive/2011/08/07/troubleshooting-ssis-package-performance-issues.aspx

Jamie Thomson - MVP

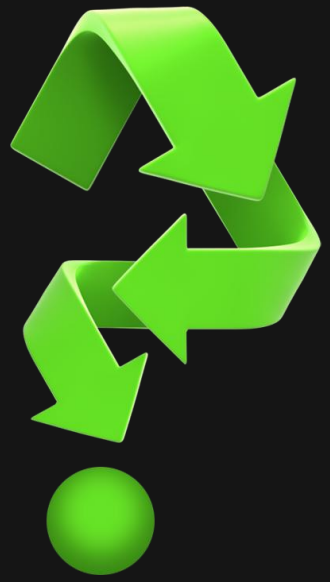
http://sqlblog.com/blogs/jamie_thomson/

http://ssisreportingpack.codeplex.com/

Julie Smith SSIS Deploy

http://sqlchicken.com/sql-university/

Questions? Need More Information?



Thank you so much! Would love feedback on presentation

Email: hope.foley@pti.net

Twitter: @hope_foley

Blog: www.hopefoley.com