Subversion


Paul van Delst

Betty Petersen Memorial Library
Technical Seminar Series
Introduction

- Branching, merging, and tagging in subversion can make parallel development of software by multiple people much easier.
  - Branching allows you to isolate changes.
  - Merging allows you control over what changes are accepted.
  - Tagging allows you to easily identify particular revisions of your code base. No more guessing!

- Goal of this seminar is to emphasise two principles:
  - Protect The Trunk
  - Minimise Conflicts

- What is your experience with branching, merging, or tagging in subversion (or other version control systems)?

- What do you want to know about branching, merging, and tagging?
Resources for this Seminar

- Version Control with Subversion (http://svnbook.red-bean.com)
- Pragmatic Version Control Using Subversion (http://pragprog.com/titles)
- Software Configuration Management Patterns (http://www.berczuk.com)

The definitive book on subversion. The website is updated frequently to reflect changes in subversion itself. We will be discussing topics from the book for Subversion v1.4 - the latest version available to us in EMC. (v1.5 is the latest stable release. V1.6 is being worked on.)
Resources for this Seminar

- Version Control with Subversion (http://svnbook.red-bean.com)
- Pragmatic Version Control Using Subversion (http://pragprog.com/titles)
- Software Configuration Management Patterns (http://www.berczuk.com)

A good introduction to the typical day-to-day usage of subversion. The book only deals with subversion v1.3, but the common usage recommendations still apply.
Resources for this Seminar

- Version Control with Subversion (http://svnbook.red-bean.com)
- Pragmatic Version Control Using Subversion (http://pragprog.com/titles)
- Software Configuration Management Patterns (http://www.berczuk.com)

Good for the pictorial representations of patterns of usage of version control while discussing the wider topic of SCM in general. Lots of “experience stories” from the author. I found I needed background/additional reading to understand the topic in the context of what we do here (i.e. scientists, not programmers, writing code).
Outline

• Branching
  – When should I create a branch? What do I call it? How do I do it?
  – What *not* to do.

• Merging
  – Trunk to branch merges.
  – Branch to trunk merges.
  – What *not* to do, common problems (well, common to me :o)

• Tagging
  – When should I create tags?

• Questions.
Pictorial “definitions”

Branch

Branch tag

Delete branch

Trunk to branch merge

Branch to trunk merge

Trunk

Committed revisions

Trunk tag

Name
Branching
When to create a branch?

• The trunk is the mainline of development - it should always pass a standard set of tests and be “nearly ready” for release or usage. As such, we always should try to **Protect The Trunk**.

• Branching is a way to isolate yourself, or others, from change.

• There are two typical scenarios to create a branch:

1. **Feature Branch**: You want to introduce a new feature into the mainline
   a. It’s not unreasonable to assume you will break the code (syntax errors, new bugs, etc) in the course of implementing and testing the new feature.
   b. You still want to be able to commit unfinished code.

2. **Release Branch**: You want to release the code
   a. Development is “frozen” for the release.
   b. Release-specific bug-fixes may be needed.
What to name the branch?

• How you name branches is up to you. Just ensure everyone involved in development knows what the convention is.

• Two examples of branch naming is EMC:

1. **CRTM**
   - Feature, or experimental, branches are named by function, e.g. **EXP-Visible** is the branch in which a visible capability is being added to the CRTM.
   - Release branches are named by their release version number, e.g. **RB-1.2**.

2. **GSI**
   - Feature branches are named after the developer implementing the feature in that branch, e.g. **mpondeca** or **dparrish2**

• The different naming conventions reflect the different development strategies each team has adopted.
How to create a branch?

- To create a branch from the trunk, you use the `svn copy` command,

  ```
  svn copy <FROM> <TO>
  ```

  where the `<FROM>` is the source URL, and the `<TO>` is the destination URL, e.g.

  ```
  svn copy https://svnemc.ncep.noaa.gov/projects/crtm/trunk \
  https://svnemc.ncep.noaa.gov/projects/crtm/branches/EXP-MyBranch
  ```

- Resist the temptation to create branches from working copies.
  - Working copies can contain mixed revisions

- A branch is simply a copy of a particular revision of the trunk filesystem.

- As I have mentioned before, there is nothing special, or branch-y, about a branch in subversion. Branches are branches only because, by common convention, we say they are.

- Branch the entire trunk. You may not need it now, but…
**Branch creation**

```sh
svn copy https://.../trunk 
  https://.../branches/EXP-TestX
```
What not to do with branching

• Preface: These are general guidelines and shouldn’t be taken as literally as written. **Context is key** - your development team should determine any strict policy-driven methodologies.

• Keep your branching as shallow as possible.
  – Branch from the trunk.
  – Try not to create branches from existing branches.

• Keep your branch lifetimes as short as possible.
  – Remember, the trunk is the mainline of development, not the branches.

• Avoid the “crawl-in-a-hole” strategy.
  – For situations where branches do exist for a long time (for a suitable definition of “long”), merge the trunk into the branch as frequently as possible.
    • E.g. CRTM project merges trunk→branches every Thursday evening. Any conflicts are dealt with on Friday.
  – Frequent updates on a branch will minimise the number of conflicts that occur at any particular merge, be it trunk→branch, or branch→trunk.
Merging
Merging cases

- We will discuss three general cases
  - Trunk → branch merges.
  - Short-lived branch → trunk merges.
  - Long-lived branch → trunk merges.

- Two different forms of the `merge` subcommand will be used,

```
svn merge -r N:M URL [WCPATH]
```

Specifies the range of revisions to merge, e.g.
- \(-r\ 1000:1050\)
- \(-r\ 1000:HEAD\)

And

```
svn merge URL1[@N] URL2[@M] [WCPATH]
```

The repository URL from which the changes will be retrieved.

The path to the working copy in which the changes will be merged. Defaults to “." (current directory).
Merging cases

- We will discuss three general cases
  - Trunk → branch merges.
  - Short-lived branch → trunk merges.
  - Long-lived branch → trunk merges.

- Two different forms of the `merge` subcommand will be used,

```
svn merge -r N:M URL [WCPATH]
```

and

```
svn merge URL1[@N] URL2[@M] [WCPATH]
```

The two repository URLs (at particular revisions) to compare for the merge, e.g.
https://.../trunk@1063
and
https://.../branches/EXP-TestX@1063

The path to the working copy in which the changes will be merged. Defaults to “.” (current directory).
Trunk to branch merges: Why?

• As previously stated, frequent updates on branches minimise the likelihood of many conflicts when the branch is merged back into the trunk.

• It also makes available to the branch any updates or bug-fixes that have been implemented in the trunk (or merged into the trunk from a different branch.)

• Each development team needs to determine the “best” frequency of regular trunk→branch updates.
  – Once a day? Once a week?
  – E.g. CRTM does weekly branch updates from trunk (using a script).

• For the more complex systems being developed here, frequent branch updates from trunk might invalidate experiments so, as always, communication between developers and code managers is key.
  – E.g. maybe a particular branch developer can “opt-out” of regular updates from trunk until their tests are completed.

• Realise that the longer the period between merges, the greater the likelihood of conflicts occurring when they are finally done.
Trunk to branch merges: Setup

1. Go to the root of your branch working copy.
2. Execute an `svn status` command,

   `svn status --show-updates`

   to determine if:
   a. There are later versions of files in the repository (* in column 7 of output)
   b. You have any local modifications (M in column 1 of output)

3. If (2a) is the case, you must issue an `svn update` command,

   `svn update`

   to bring your working copy up-to-date.

4. If (2b) is the case, you should commit your local modifications prior to performing the merge,

   `svn commit`

5. At this point you have a “clean working copy”
Trunk to branch merges: First Time

1. Determine the trunk revision from which the branch was created; let's say it was 1000. Add one to give 1001.

2. Determine the current trunk revision in the repository; let's say it is 1050. Subversion also recognises the keyword `HEAD` for this case.

3. Issue the `merge` subcommand with the `--dry-run` switch,

   ```
   svn merge --dry-run -r 1001:1050 https://.../trunk.
   ```

   This will list all the changes that will occur, without actually doing anything to your branch working copy so you can see if there are any conflicts.

4. If there are no conflicts, or their number is reasonable (more on that later), reissue the `merge` subcommand without the `--dry-run` switch to actually perform the merge in your branch working copy.

5. Deal with any files in conflict and resolve them,

   ```
   svn resolved <FILE>
   ```

6. Commit the merge changes with a useful log message,

   ```
   svn commit -m "EXP-TestX branch. Merged trunk r1001:1050 into branch"
   ```
Trunk to branch merges: First Time

```
svn merge -r 1001:1050 https://.../trunk.
```

```
svn commit -m "EXP-TestX branch. Merged trunk r1001:1050 into branch"
```
**Trunk to branch merges: Next time**

1. Determine the end revision of the last trunk merge into the branch by looking at the log message for the branch,

   ```
   svn log | more
   ```

   In our example that was 1050.

2. Follow the same procedure as for the first time.

   - Why the separate slide? *To emphasise the importance of specifying the merged revisions in the commit log message.*

   - Without manually tracking the merged revisions in the commit log message, you have *no way* to determine which revisions from the trunk have been merged! (Subversion v1.5 does it for you, but…)

   - If you remerge previously merged revisions, you will typically get many, many conflicts. If this happens, it’s a clue that the merge revision range is probably incorrect.
Using trac SCM tool to inspect commit logs
Branch to trunk merges

- There are two basic scenarios to deal with in merges branches back to the trunk.

1. The branch is short-lived and has had no updates from the trunk.

2. The branch is long-lived and has had several updates from the trunk.

- As with the trunk→branch merges, you’ll want to merge into a clean working copy of the trunk.
Subversion: Branching, Merging, and Tagging

Short-lived branch to trunk merge

```
svn merge -r 1001:1063 https://.../branches/EXP-TestX .
```

```
svn commit -m "Merged EXP-TestX branch r1001:1063 into trunk"
```

```
svn delete https://.../branches/EXP-TestX
```
Long-lived branch to trunk merge (1)

- Branch required multiple trunk→branch merges, committed at revisions 1016 and 1051.

- We now want to merge the branch back into the trunk so trunk developers have access to the branch updates. Note that branch developer may still want to work on further branch changes.

- What merge range in branch? How to avoid remerging the trunk updates?
• Perform a “final” trunk→branch merge and commit.

```bash
svn merge -r 1051:1071 https://.../trunk .
svn commit -m "EXP-TestX branch. Merged trunk r1051:1071 into branch"
```

• If we now difference the HEAD (or revision 1072) of the branch and trunk, the result will be only those changes made in the branch.
• Now we switch to a clean trunk working copy and use the second form of the `merge` subcommand

```bash
svn merge https://.../trunk@1072 https://.../branches/EXP-TestX@1072 .
svn commit -m "Merge of trunk@1072 and EXP-TestX@1072 to trunk"
```
Using trac SCM tool to inspect commit logs

**CRTM**
Community Radiative Transfer Model

**root / trunk**

<table>
<thead>
<tr>
<th>Rev</th>
<th>Chgset</th>
<th>Date</th>
<th>Author</th>
<th>Log Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>5921</td>
<td>[5921]</td>
<td>23 hours</td>
<td><a href="mailto:yang_hen@noaa.gov">yang_hen@noaa.gov</a></td>
<td>trunk src: Zeeman subdirectory * Zeeman Utility (load_bias_field.dat) ...</td>
</tr>
<tr>
<td>5903</td>
<td>[5903]</td>
<td>28 hours</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>test StfOptics/RSSEM subdirectory. * Makefile. Updated SRC_FILES ...</td>
</tr>
<tr>
<td>5902</td>
<td>[5902]</td>
<td>28 hours</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>test StfOptics/RSSEM subdirectory. * Test, RSSEM.PRD: Updated ...</td>
</tr>
<tr>
<td>5854</td>
<td>[5854]</td>
<td>7 days</td>
<td><a href="mailto:paul.vandela@noaa.gov">paul.vandela@noaa.gov</a></td>
<td>Merge of trunk@5853 and EXP-Zeeman@5853 to trunk</td>
</tr>
<tr>
<td>5816</td>
<td>[5816]</td>
<td>8 days</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>test StfOptics subdirectory * RSSEM: Added new test subdirectory ...</td>
</tr>
<tr>
<td>5792</td>
<td>[5792]</td>
<td>12 days</td>
<td><a href="mailto:yang_hen@noaa.gov">yang_hen@noaa.gov</a></td>
<td>Fix TauCoeff/ODR/Infrared/ODR subdirectory. * ...</td>
</tr>
<tr>
<td>5791</td>
<td>[5791]</td>
<td>12 days</td>
<td><a href="mailto:yang_hen@noaa.gov">yang_hen@noaa.gov</a></td>
<td>Fix TauCoeff/ODR/Infrared/ODR subdirectory. * ...</td>
</tr>
<tr>
<td>5788</td>
<td>[5788]</td>
<td>12 days</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>test Coefficients/TauCoeff/Test_TauCoeff subdirectory. * Makefile. ...</td>
</tr>
<tr>
<td>5787</td>
<td>[5787]</td>
<td>13 days</td>
<td><a href="mailto:paul.vandela@noaa.gov">paul.vandela@noaa.gov</a></td>
<td>scripts/ruby/svn_util/00 * merge_trunk2branches.rb: initial commit.</td>
</tr>
<tr>
<td>5781</td>
<td>[5781]</td>
<td>13 days</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>test Coefficients/TauCoeff/Test_CRTM_TauCoeff subdirectory. * ...</td>
</tr>
<tr>
<td>5780</td>
<td>[5780]</td>
<td>13 days</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>src/Coefficients/TauCoeff subdirectory. * TauCoeff_Define.f90 ...</td>
</tr>
<tr>
<td>5779</td>
<td>[5779]</td>
<td>13 days</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>src/Coefficients/TauCoeff subdirectory. * TauCoeff_Define.f90 ...</td>
</tr>
<tr>
<td>5778</td>
<td>[5778]</td>
<td>13 days</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>test LifeCycle subdirectory. * Makefile: Updated SRC_FILES to ...</td>
</tr>
<tr>
<td>5777</td>
<td>[5777]</td>
<td>13 days</td>
<td><a href="mailto:david.graff@noaa.gov">david.graff@noaa.gov</a></td>
<td>test Coefficients/TauCoeff/Test_CRTM_TauCoeff subdirectory. * ...</td>
</tr>
</tbody>
</table>
Using trac SCM tool to inspect merge result
What not to do with merging

• Remerge previously merged changes.
  – This has been mentioned several times.
  – Due to the version of Subversion we have to use (v1.4), we need to be responsible for keeping track of merges between the trunk and branches.
  – Subversion v1.5 solves this problem by basically attaching the merge information as a property of the file.

• Merge into a dirty copy.
  – Cannot separate local modifications from merge operation.

• Incorrectly resolving conflicts.
  – This can result in branch, or trunk, changes disappearing from a merge operation. “What!?” I hear you say?
  – But, easy to rectify if the merge is done into a clean working copy.

• Merge across branches.
  – Things can get messy really quick.
Merge into a dirty copy

These files were all local modifications. I can no longer separate the commit of the merged files from the commit of the local mods.
Incorrectly resolving conflicts (1)

- About a month ago, I did a “routine” trunk→EXP-Visible branch merge in the CRTM.
- There was a conflict in a file that I resolved by simply replacing the working copy with the latest revision.
  - Unfortunately, the latest revision was a trunk copy so all the changes made in that file in the branch “disappeared” in the branch HEAD revision.
  - Needless to say the branch developer, Quanhua Liu, was a bit confused the next day when he sat down to continue working on the branch.
- How to fix it?

1. *Undo the merge operation* for the file in the branch working copy (and commit!)
   
   ```
   svn merge -r 5622:5621 CRTM_RTSolution.f90
   ```

   This was doable because the initial merge was done into a clean working copy.

2. *Remerge* the trunk into the branch just for this file to bring it back into conflict
   
   ```
   svn merge -r 5518:5621 \ 
   https://.../trunk/src/RTSolution/CRTM_RTSolution.f90
   ```

3. *Correctly resolve the conflict* and commit (with a suitably humble commit log message)
Incorrectly resolving conflicts (2)
Incorrectly resolving conflicts (3)
Incorrectly resolving conflicts (4)

Changeset 5633 for branches/EXP-Visible

Timestamp: 11/09/09 09:55:18 (3 weeks ago)
Author: paul.vandela@noaa.gov
Message: EXP-Visible branch.

src/RTSolution
* CRTM_RTSolution.f90: Remerged trunk r5518:5621 to correctly merge this file. Replaced microwave sensor id with that for infrared.
(CRTM_Compute_RTSolution): Correctly merged only that part of the trunk dealing with the trunk changes made in r5621.

Files: 1 modified

branches/EXP-Visible/src/RTSolution/CRTM_RTSolution.f90 (3 diffs)

Done
This may be a bit of a red herring as I encountered this problem merging back when I had no idea what I was doing.

I had done the following:

I now wanted to merge all the trunk changes back into the **EXP-SOI** branch.

How to do it without remerging that section of the **RB-1.1** branch that made it into the **EXP-SOI** branch in **r1866** back into the branch after the **RB-1.1**→trunk merge?

At this time I was unaware of the second form of the **merge** subcommand we’ve looked at, so my solution may be overly contrived due to my ignorance....
Merging across branches (2)

- My solution...

- The moral of the story: If you can help it, don’t do cross-branch merges. Always try to merge to/from the trunk.
Tagging
When to create a tag?

- A tag is simply a snapshot of a particular revision of the repository.

- It's basically a way to give a human-friendly name to a particular revision rather than having to remember a revision number.

- You typically create tags:
  - At the beginning and end of short-lived development branches for easy merging.
  - For releases; alpha, beta, and final.
  - Before, and after, complicated merge procedures (just in case)
  - If you have a particular combination of revisions in your working copy, and want to create a tag of that (I don’t do this, but it can be useful to folks)

- Anytime you want to refer to a particular trunk or branch revision in the repository, for whatever reason, create a tag (with a meaningful name, of course).
How to create a tag?

- A tag is created in exactly the same way as a branch, using the `copy` subcommand, but you copy it to the tags repository directory.

- Let’s say I tag the beginning of my **EXP-TestX** branch, before I’ve made any changes in the branch,

```
svn copy https://.../branches/EXP-TestX \ 
  https://.../tags/PRE-TestX
```
So what?

- O.k., I’ve created a tag of the start of my branch. So what?

- Well, let’s also tag the end of my **EXP-TestX** branch, after I’ve made all my changes and want to merge to trunk,

  ```
  svn copy https://.../branches/EXP-TestX \
  https://.../tags/POST-TestX
  ```
Use tags for merging

- Now let’s merge the branch changes to the trunk,

```bash
svn merge https://.../tags/PRE-TestX \ 
  https://.../tags/POST-TestX .
svn commit -m "Merged PRE-TestX to POST-TestX tags into trunk"
```

- So, you can use tags to avoid mucking about with revision numbers.
- This works best with very short-lived branches (e.g. bug-fixes). In these cases you might use a bug tracking number in the tag name.
Use tags for releases

• You can tag various stages of your release branch as you iron out problems.

• Let’s regularly tag a RB-2.0 release branch:
Use tags to bracket merges

- This usage is a defense mechanism to protect the trunk.

- If anything bad/strange/whatever happens with a merge to trunk, you can always easily restore from a tag.
Summary
Summary

- *Protect the trunk*
- *Commit often*

- Branches
  - Keep them shallow
  - Create them from repository, not working copies

- Merging
  - Merge as regularly as your development strategy can handle
  - Track your merge revision numbers (SCM tools like trac are invaluable)
  - Merge into clean working copies only
  - Be careful resolving conflicted files (but don’t obsess)
  - Avoid cross-branch merging if you can help it.

- Tagging
  - Tag with abandon!

- Subversion copies are cheap operations so don’t hesitate to branch or tag.
Future Library Technical Seminars

- If you want to request a topic for a Technical Seminar, or want to volunteer to give one, contact Jan Thomas. See the library “Ask the Librarian” webpage:
Questions?