1 - Master Documents

This chapter ventures deeply into Microsoft heresy. A heretic is someone who preaches heterodoxy, or mixed doctrines. Unlike a lot of official MS and MVP speak, this topic advocates the usage of a certain feature that can be said to be generally considered as broken - Master Documents, or Masters. As so little information is forthcoming on this subject from other sources, yet many writers use them regularly because there is no other choice, it is fully covered here.

It provides a complete explanation of Master Documents and how to use them. This chapter helps you:

- Minimize their potential for trouble
- Minimize your efforts to produce large documents or easily re-use source material
- Enable extremely large document handling with minimal risk
- Collect multiple documents into a single, consecutive sequence of pages with simple cross-referencing.
- Concentrate on writing rather than document management in complex document development situations.
- Help control situations where tools, such as document converters, demand situations where multiple slave files are required to a master.

As Masters and large documents have many points in common, there are also some useful tips and tricks for dealing with common large document problems.

Definitions

Before exploring the usage of a Master Document, Master from here on, some background information is needed on the nature of the beast that lurks within.

As usual, the interface is referred to as the façade, a much more suitable term in most cases, although less so in this situation. Most of the background functionality is made available through the user interface. Some of it is downright useless; some of it is worth using.

What is a Master?

A Master is able to contain a number of Sub-Documents, or Subs. This is exactly the same hierarchy as exists between files and folders – a Master contains Subs just as a folder contains files. The Subs exist as files in their own right, and indeed are treated as such for most of their life. These Subs may have to exist for a number of reasons, but at the end of the day all the Master is is a skeletal container for a number of other Word documents. The manner in which this is implemented is, of course, a huge kludge.

However, treated with respect, Masters can not only produce absolutely enormous documents - so large that if the limits were exceeded, the document produced would not be usable as a single bound volume anyway – but also as the ultimate publishing binder for collecting separate documents. The Merriam-Webster's 3rd International Dictionary is an example of the upper practical limits for printed works.

Why Not to Use Master Documents

Master documents excel at generating corruption, corruption that is totally hidden and can completely destroy all associated documents. Additionally they are more work to maintain than a single small file; rebuilds must occur more regularly and the work of inserting and maintaining more files. This is compared to small files only, against large files or collections of separate documents, time and effort savings are substantial.

Note also that ALL word documents – whether Sub, Master or stand-alone – are heading towards total corruption anyway. Good working practices involve regular rebuilds when the revision number is starting to climb into the hundreds, so there isn't much extra to do to work with a Master.

There is a lot less work required to maintain a document set and Masters are less prone to errors than the alternatives such as RD fields with hard-set page numbering.

If a document is going to be under 500 standard pages, you shouldn't have to resort to a Master. Of course, there are many more reasons than just size alone to keep documentation in separate files yet require binding together at publishing time.

If a document will grow to be larger than this, whilst it is small and weedy it is better off as a single document. Split it into chunks and use the Master when it starts getting really big.

Masters still corrupt slightly faster than single documents when the document size is enormous. They also have the potential, through sloppy or naïve (that is, normal) work practices – especially saving the changes after every print and editing the Subs via the Master – to corrupt faster and more thoroughly than any other document type.

Why Use Master Documents?

A few very good reasons, and when you find yourself with this requirement, you can implement a simple solution.

Big Files

Lets face it; Word has an upper limit on file size. Sooner or later you'll run into it. It gets chewed up as noted below in File Sizes. Big Files aren't just big with bloat – all the formatting and layout has to be stored somewhere. Rich, structured documents running past 500 pages start to enter the Big Files category.

Big Files have more potential corruption containers and methods so they corrupt faster. Thus, keeping work files small reduces the overall corruption of each document.

Additionally, sometimes network performance can slow file access to large files to the point where productivity is regularly being lost — especially for highly dynamic manuals.

This also means you are much better off operating on these files in a local environment, not across a network. Rule 10 states that version control software helps, one of the hidden benefits is it is very easy to set the version control software up so that it maintains copies of the files on your local disks.

File Sizes

The clearly defined limit for a Word file size is 32 Mb and this excludes all graphics. The number of pages in a word document is also a bad indicator, as the more included screen-resolution pictures blows out the possible number of pages which is limited more by applied styles and textual/structural content. A structurally rich file contains less content so the reverse side of the coin is using styles and avoiding all manual formatting leads to greater document stability.

Editioning

Master documents can be used to implement simple editioning to aid in single-sourcing common material amongst a document set. You might have substantial amounts of text that are always included in certain documents. These chunks can be saved as their own file and just inserted where required.

In simple terms:

- 1. Identify large chunks of repeated information and turn these into separate Subs.
- 2. Turn the remainder of the newly mangled document into separate documents, one for each chunk and the bits in-between. For example a 10-chapter document shares chapters 4,5 and 8 with another. The minimum Subs that must be produced are the chapter ranges 1–3, 4–5, 6–7, 8 and 9–10.
 - With smaller chunks than chapters, it can seem a bit more confusing. That's why the practical limit to this practice is where paragraphs differ. It can get very messy maintaining and editing tiny Subs.
- 3. Build master's to contain the chunks. To continue the example above:
 - Master 1 has Subs Ch1-3(Master 1), Ch4-5, Ch6-7(Master 1), Ch8 and Ch9-10(Master 1).
 - Master 2 has Subs Ch1-3(Master 2), Ch4-5, Ch6-7(Master 2), Ch8 and Ch9-10(Master 2).

This technique, when combined with the Editioning dialog discussed later in Utilities, can help extend other single-sourcing capabilities for Word.

As an example of how this can work even small chunks of information, imagine we have just one substantial document chunk that belong sin the preface of most documents we produce. It has the contact information for the company and the publications department, the document conventions, some legal requirements and other common front matter.

A typical published document in this scenario consists of four files:

- 1. The beginning of the preface with a unique title page and introduction
- The standard included matter.
- 3. The rest of the published document.
- 4. The Master to contain it all.

Requirement to have Separate Files

Whether due to geophysical, political, design or other reasons, sometimes it is just not possible to combine all the files into just one. Master documents can save a lot of effort by referencing all the source documents in their respective locations for ease of

presentation and publishing. The writer does not need to overly worry about sequential page numbering, consistent styling or cross-referencing between the individual documents.

Masters support both relative and absolute paths. Material that is under your control can be stored relative to the master. Material that isn't can be linked absolutely. Most of the time if you cannot have control over where the latest material is presented it has to appear in a static location anyway.

Built-in Tools

All of the tools for maintaining Masters can be found in the façade on the one toolbar. In Word 2000 and later this is View > Outline, in earlier versions View > Master Document. It is one of the few features where the non-VBA user has full access to all the functionality.

As the façade does provide the power to drive this feature, you need to be familiar with the façade's toolset to use Masters with minimum effort. This section helps familiarize you with the basic toolset you use to create and control Masters and Subs.

There are three groups of controls on the toolbar. The first – the arrows, promote and demote heading levels, the next are the numbers that set the current depth of expansion and the last set are the focus of this topic – the Master document toolset.

Expand/Collapse Sub-documents

This collapses or expands all the Subs. Collapsing a Sub is similar to collapsing a heading using the map or outline view of a document. It causes it to close and visually remove itself from the façade. Thus collapsing all the Subs leaves you with a view of the document structure. The initial view of a loaded Master is with all Subs collapsed.

Expanding all Subs is the opposite. All Subs are opened and their content is loaded into the Master. This is used mostly at production and proofing time, as well as expanding the Subs to perform cross-referencing across them as per Rule 2.

Create Sub-document

Unused if you follow the Heretical rules. This takes the highlighted range and creates a Sub out of it. The Sub is not saved to disk explicitly by this process and acquires a VBA property of ActiveDocument.IsSubDocument of True. This is the only way, other than the equivalent VBA command, that a document gets this property value.

It isn't used in the Heretical rules because you don't know exactly what craziness Word decides to attach to these files. You get far more control, as well as a simpler, easier to understand process for generating the files, by using standard cut and paste on individual documents to generate your document set.

Remove Sub-document

Unused if you follow the Heretical rules. It does not remove the content – instead it removes the boundaries of the selected Sub so that the content is now part of the Master.

To actually delete a Sub from out of a Master, select the range that it belongs to and delete that. This still doesn't delete the disk file; it just removes that Sub from being linked into the Master.

This feature is not recommended as it forces the Master to contain text, which even if deleted can leave undesirable corruption containers lost in Master space.

Insert Sub-document

You use this function to place your Subs into your Master. It opens a file browser to select the file to insert into the current document as a Sub – thereby turning the active document into a Master and the selected document into a Sub. This change, a document becoming a Master or a Sub, is not as radical as it may seem. Very little happens to either document other than the Master obtaining a Subdocument object in the required range.

This is the main tool to use; if you use Masters regularly you may want to place this single control on an exposed toolbar as there is no way around using this function if you want o use Masters.

However, the functionality is slightly flawed. The Master's current path is not used as the default path for this dialog, yet the default path is used as the offset for relative addressing, meaning that absolute references are generated unless you take a simple precaution.

Before adding the first Sub in an editing session, use the Insert Sub-document dialog to browse to the directory where the Master is. Select OK without selecting any file. This will reset the default path and cause relative addresses to be generated when you insert the Subs.

Merge Sub-document

Unused if you follow the Heretical rules. Turns two or more Subs and the intervening space into one. Although not nearly as dangerous as some of the other tools, it is far better to explicitly create and destroy your documents using the file system, rather than relying on Microsoft's attempt to get it right via Word's 'logic'. Additionally, any formatting discrepancies will be left at the mercy of Word's marvelous foresight to correct. I always avoid anything that makes Word think – it invariably gets it wrong.

Split Sub-document

Unused if you follow the Heretical rules. Turns the Sub into two at the selection point. The new document will require a name when saving. You may be tempted to use this when you realize a topic requires sub-dividing. Resist the urge, open the Sub by itself and cut out the text required for the next document and paste it into a new document by hand.

Lock document

Prevents changes to that Sub. Subs are automatically locked if the author name is not the same as the current user name. In outline view, a locked document has a small padlock on its top left corner. Locking a document is an excellent way of preventing changes to that document whilst you spend your few perilous moments in Master mode,

but if you follow the rules, there is no need to use this function as it just becomes a burden on the user.

The flipside to this practice requires the author's name to be set to the user id that performs the publishing of the documents to reduce complications like the page numbers in a locked document not updating.

VBA

Whilst there is little to do directly with Masters directly in VBA other than rebuild them, there are macros to aid in the assembling of large documents. Rebuilding of documents is not within the scope of this topic.

The Document Object Model

There isn't a whole lot hidden behind the façade in this case, that can be revealed by diving into the VBE. The first two most obvious settings of interest to the VBA programmer, are ActiveDocument.IsMasterDocument and ActiveDocument.IsSubdocument.Both of these are Boolean values. The IsSubdocument value only indicates that the document was created with the Create Sub-document method or button. It does not mean that the document belongs to a Master document as a Sub.

The Subs themselves are referenced through the ActiveDocument. Subdocuments collection. This has several properties.

- HasFile (Boolean) indicates that the Sub has been saved as a file. Of interest only when using the not recommended tools.
- Level (Long) indicates the heading level used to create the Sub. If the Sub wasn't created with the Master document toolset, which is the case if you are following this document, the number is set to 1 and is meaningless.
- Locked (Boolean) indicates whether the Sub is locked for edit or not.
- Name (String) is the Subs file name.
- Path (String) is the Subs path.
- Range (Range) is the range (a pair of numbers the start and end points) in the document that the Sub covers.

There are three methods for a Sub - Open, Delete and Split that perform as per the façade's toolset. Note that Delete does not delete the disk file nor remove the contents of the Sub; it just removes the Sub's boundaries and existence in the Master.

Custom

Linking headers and footers is a common activity. This macro cycles through every section in the document other than the first, using its page set up as a guide to which headers and footers exist in the section and then links them all to the previous.

Macro: Link Header and Footers

```
Public Sub LinkHnF()
Dim k As Long
With ActiveDocument
If .Sections.Count > 1 Then
```

Other useful macros, described elsewhere, include:

- The Page Numbering tool, for tweaking up different page numbering styles for each section for publishing.
- The Rebuild tool, for rebuilding your documents.
- The Style Gallery tool, for enforcing presentation consistency.
- The MegaXRef tool, for inserting complex cross-referencing.

Using Masters

The simple explanation is to operate Masters in a clean room environment. Isolate them from the precious text and formatting. Merely use them as a linking reference for publishing purposes. Many people who read this topic walk away saying to themselves "Too much work". Rule 8 – embodied in a work process means there is no extra work to do! At the end of the day, when documents are that close to production they have already been substantially worked over. They require a rebuild anyway. Thus a rebuild that includes creating the Master or reusing the standard Master, and quickly producing the production copy is almost no extra time and effort.

The Ten Heretical Rules of Masters

Following the simple rules below drastically reduces problems with using Masters to the point where they are a viable mechanism for large or complex document production. These rules have been tested on documents exceeding a thousand pages with over fifty Sub-documents (Subs). The resulting documents were too large to be a single file, or even two. Sometimes the documents were authored or revised on multiple machines by multiple users.

A point of note is that with so many forced empty paragraphs to wrap the Subs in, the topic heading styles, used to start the Subs, should not have any Space Before set.

Rule 1 - No Text in the Master Other than the Auto-TOC

Even the title page and preface material need to be in Subs. Ideally each Chapter gets its own Sub. Rule 2 aids this rule, as it is possible to accidentally insert text between Sub boundary markers and have it isolated in the Master.

This is the primary rule as this is the text that invariably corrupts first and can cause the Master to hold multiple conflicting style instances, ensuring the early demise of the document set. As this is not a desirably outcome, using this rule alone substantially reduces all efforts to maintain a Master structure.

The rule bends to allow one or two heading levels to appear in the Master as text. This is often quite handy to get an outline view of the project without having to expand the Master. To aid in rebuilds, duplicate these headings into a standard document for copying into new Masters.

A sub-rule to this is to not use Masters that are in turn Subs to another Master. Not only does the primary rule above turn this into an unnecessarily deep hierarchy, but also because of inheritance, that structure is enforcing cascading corruption. Not a recommended result for anyone's books.

Rule 2 - No Editing Sub-documents via the Master

Do NOT perform editing via expanded Subs from the Master. Although not as bad, also avoid editing via double-clicking the Sub from within the Master. Edit the Subs directly from the File > Open dialog. Doing this minimizes the Master's interaction with the document and reduces the speed and spread of corruption, which in turn reduces overall downtime for recovery and thus administration overheads drop.

Opening a document that is linked to others, for example by Sub or RD fields, by any other method than using Word's File > Open is dangerous. Many documents use relative path addressing and the only way to set the current path correctly to resolve these filenames is to use the Word's File menu option Open. So avoid opening a Sub via its Master, if you do try to avoid linking to files external to the Master's content.

The subtle side of this rule is to not save the changes after having the Master open for a print run as is re-emphasized by Rule 9.

The two inescapable exceptions are described separately. Adding cross-referencing is described in Rule 6 and removing the auto-section-breaks in Rule 3.

Rule 3 - Delete all auto-Section Breaks

Sections, like Masters themselves, are an excellent source of corruption. Minimizing their number minimizes corruption. Additionally, as most Masters have consistent styling and layout throughout, they make it harder to enforce consistency. Use the same template in both the Master and the Subs to improve this further, as described in Rule 4.

However, as also described in Rule 4, there is one situation where the automatic section break should be kept. It is for its original purpose, which is isolating the formatting in that section, for the situations where you have common style names but different format definitions.

The easy way to delete all section breaks is using the FnR dialog. Find ^b and replace all with nothing.

This rule has several practical implications:

- After inserting a Sub, immediately delete the section breaks at the start and end of it.
- As the final section break attaches itself to the last paragraph mark in the Sub to act as the corruption container for it in the Master, every Sub must have a blank paragraph to end it before adding it into the Master.
- With the section breaks gone the immediate visual aid to the boundaries of the Sub is gone. The vague outlined boxes are inaccurate boundary markers. To reduce the chances of merging or interleaving Sub and in addition to leaving at least one blank

paragraph at the end of each Sub, leave a blank paragraph between Subs. This also simplifies inserting a Sub between two others later on as it provides a space in the Master to select and press **ENTER** to form a blank paragraph pair in the Master to insert the Sub safely between.

· Consistency is maintained with less effort.

If you must have section breaks in your document, any document for that matter – not just Masters – only place these breaks as part of your final publishing run. This reduces the opportunity for corruption and reduces maintenance effort on the document.

Rule 4 - Master and Subs to have the Same Template

This not only aids in presentation consistency but also helps reduce the decisions that Word must make to resolve or compile a Master document, which means less corruption and a reduced frequency of strange results. The next Rule describes the decisions in more detail.

Using this rule means not only fewer opportunities for corruption to occur as well as simplifying your infrastructure requirements.

However, when Masters are being used to collect together documents, occasionally the style set in the target documents is completely different. This is a difficult situation. If any styles have the same names yet different formatting from the ones you use, you cannot rename them nor keep them entire. Section breaks in this case help maintain the separation in definitions, but do not try and rename the styles as you do not have control over that source document.

Rule 5 - Rebuild the documents when the Template style changes

Rebuilding documents is described in the chapter on styles. Rebuilding them helps prevent the corruption that can occur when overwriting styling information for every paragraph. It also helps to reduce inconsistencies in the final document.

Depending on document settings, new styles from the template may overwrite the Sub styles. Once the document is integrated into the master, the Master's definitions may overwrite the Sub styles. Every paragraph that has its style overwritten can cause a little bit of corruption.

For workflows that are going to have regular style changes during development, rebuilds can be left for late in the production cycle. Also note the tie-in with Rule 8. This rule can be bent as documents can carry a fair amount of corruption before functionality is impacted, but at some point you have to follow it to reduce your built-up corruption or unimplemented updates.

Careful planning sees the use of Rule 8 come into its own. All the 'dangerous' activities – creating the Master, removing the section breaks caused by the creation and adding cross-references between documents – are performed at the end of the development process. Overall, this reduces downtime from the infinite array of Word corruption errors.

Rule 6 - Ignore Rule 2 to Cross-Reference between Subs

This is the only activity that requires all Subs be open in the Master and editing to take place. With the automatic bookmark generator for headings in the cross-references

chapter or with using a consistent naming scheme it is possible to insert the crossreference field manually and entering the appropriate bookmark name by hand to avoid this requirement, however this is usually more work than it is worth.

Keeping the advice from the cross-referencing topic, use bookmarks to set your destinations rather than Word's heading numbers or list item numbers. Hyperlinks get used to reference external documents only. Ensure they have a fully qualified URL. Cross-references are used for internal links.

For very large Subs with many 'external' cross-references it is also advisable to perform a simple rebuild of the Subs and regression of the Master after completing the cross-referencing to throw out any corruption that may have gathered during this process.

This rule ensures excessive effort is not required to maintain reasonably flexible documents and reduces the production time for linking multiple documents together in an interactive fashion.

Rule 7 - Regularly throw the Master out and Rebuild it from Scratch

This is why Rule 1 says to keep the TOC in the Master; print it out without updating to use as a rebuild guide. This is also why implementing Rule 8 is advisable.

To make the "perfect Master" follow the Creating a Master instructions below.

Corruption tends to build up in the Master itself. It is totally invisible to you and ultimately lethal. Thus, the best way to deal with this unseen enemy is to discard their entire territory and start afresh, thereby reducing your overall downtime for recovering from corruption.

Rebuilding should occur after the Master has been updated a few dozen times. By this time the Master has picked up undesirable formatting conflicts and probably excessive style and list definitions.

Rule 8 - Only create and load the Master for publishing

Do reviewers really need to see all consecutive page numbers between sections or completed cross-reference links? Not until the final stages of the development process. Before then content is still king and formatting is local to each document. There is some slight hassle to having three or four documents open at once rather than one, but even this only establishes itself well into the development of the work.

Maintenance of dynamically published, or live, manuals is even easier. The change is affected to a topic in a section or the page reference can be manually converted to such. The Sub is loaded, edited and saved. The standard publishing instructions described below apply and nothing is affected. A common practice to minimize printing requirements is to restart page numbers at each chapter breaks. That way only the chapters that contain changes need be printed.

Sometimes it is just not possible to leave the Master that late in the development process. Rule 7 contains the brief instructions to building the Master and the Creating a Master procedure below contains more detail; use copies of this Master to generate review copies. Try and hold off on this as long as possible so that the structure has stabilized to avoid unnecessary rebuilds or edits. Avoiding building a Master until as late in the development process as possible reduces your overall effort and administration costs.

Additionally, come publish time, the Subs need to be unlocked. The easiest way to achieve this is to set the author's name on each document to be the same as the user who performs the publishing.

Rule 9 - Do NOT save changes after publishing

That is – develop the Subs to perfection first and forget the updated links and so on from the print process. Unwanted corruption can get saved back into the files. Version control software can eliminate this requirement to a point, by regressing or ignoring the changes, by using instructions similar to those for building the Master described in Rule 10.

Once you press Print and send the Master off to publish heaven, it has had its fields updated and wants to save the changes. Don't. This is another rule aimed at reducing downtime recovering from corruption, yu just deny the opportunity for this to happen.

Rule 10 - Use version control software

In practice, this can be simulated to a small degree with a good old-fashioned copy of the working directory. The idea is to get around the Master's tampering with the Subs.

- 1. Get the Subs ready for insertion.
- 2. If there is no version control software then delete any old Master document and backup the directory. Otherwise check all the files out.
- 3. Build the Master.
- 4. If there is no version control software then copy back the directory. Otherwise check the Master in and undo the checkout on all the Subs. Avoid checking the Master back in ever again.

Of course, version control software allows regression from corrupt states whereas the backups need to be performed regularly and be retained for long periods to use them to recover from corruption. In such rare cases as regression is required, don't forget that even though the document appears perfect it has a school of corruption sharks lurking under the surface waiting to come and bite you. Rebuild the document thoroughly at once.

This is why multiple development versions are preferred. You may need to regress your document several versions on extremely bad occasions. These methods reduce the risk to the point where you shouldn't be affected, but it is always better to be safe than sorry. Following this rule ensures you will never lose too much work to radically your department.

Finally, a pleasant side effect of this rule is you can set the version control software to keep your copies of the files on your local disks, reducing errors, conflicts, lag, outages and false failures.

Creating a Master

An overview of a quick and easy way to produce very large documents, or documents consisting of many separate word files with zero hassles. Word files can be used to embed many other data types, so this procedure can gather many types of information for publishing in a single physical document.

Considerations

Use Version Control Software, such as VSS, to store changes in. This gives complete version control over documents, and provides a simple way to regress corruption should it arise.

The authors name must be set to yours in each Sub to be able to work on the Master with its Subs.

Prerequisites

Create your template and get it right. Start mucking with templates and styles after a Master has been set up and you are begging for corruption. Use the same template, where possible, in both the Master and Subs.

The document is to be over 500 A4 pages, to have multiple authors working on it simultaneously or exists as a collection of separate documents for development, political, physical or other reasons. Otherwise it's simply not worth the extra effort and risk.

A set of documents ready to be compiled into one larger document.

Procedure

- 1. Make a copy of all the files to be Sub-documents.
- 2. Create a document and save it with the proper name for the final work. This will be your Master.
- 3. Use the Insert Sub-document feature to browse to the path the Master is in to set the relative path base. Select OK.
- 4. Use the Insert Sub-document feature to insert the Subs in order:
 - Leave a blank line in between each Sub. This avoids a lot of screwing around to insert a new Sub in between two others without it becoming a Sub of a Sub.
 - When asked for "Rename Styles?" do NOT rename any. If you choose the wrong option, hit undo after it loads in and redo it correctly. Then insert your table of contents with its heading after the preface and before chapter 2.
 - Delete all section breaks. If you must have a heading on an odd page, add the section breaks as part of the proofing run, which is the last very stage of document delivery.
- 5. Other than the TOC and its heading, no other text should be kept in the body of the Master document.
- 6. Set your header and footers inside the Master another good reason to always do this via macros in the template. The header/footer are stored as Autotext and get placed by a simple macro. Ensure all headers and footers are linked through if you have multiple sections.
- 7. Ensure Tools>Options>Print has Update Links ticked. Do a **CTRL+A** then **F9** (select all, update fields) to ensure the latest information is present.
- 8. Exit the Master.
- 9. Copy all your archived Subs back over the top of the active Subs and the procedure is complete.

Troubleshooting

If you can't avoid template changes:

- 1. Develop your Subs as separate documents based on one template.
- 2. Fiddle with this template shamelessly as you work.
- 3. When it comes time to publish, ensure your style editing has stabilized for the proofing run.
- 4. Rebuild the Subs via cuttenpaste-all-but-the-last-paragraph-mark and use update styles from the Tools > Templates and Add-ins dialog. You would have had to do this anyway because of template changes.
- 5. Build a new Master and use it to publish.