Who do contents of a file get corrupted when I use EFT to offload a text file with Unicode contents to another server?

THE INFORMATION IN THIS ARTICLE APPLIES TO:

- EFT Server, v6.x and later

QUESTION

Who do contents of a file get corrupted when I use EFT to offload a text file with Unicode contents to another server?

ANSWER

When transferring files in ASCII mode (TYPE A), EFT will examine the file for Line Feeds (LFs) and insert Carriage Returns (CRs) before line feeds if those are missing, in accordance with RFC 959 §3.1.1.1. Text files with Unicode contents saved in Unicode or Unicode big endian format will contain extra NULL characters between the CR and LF sequences, and may contain the same hex value as LF (0x0a) in place of other Unicode characters throughout the file's contents. EFT Server detects those isolated LFs or those LFs with preceding NULL characters and inserts CRs in front of each one, resulting in corrupted file contents and line endings.

The workaround is to choose UTF-8 as the encoding rather than Unicode when saving the file. This will preserve the contents and the correct CRLF sequence resulting in zero changes by EFT prior to offload.

Another solution is to force EFT to transfer the file in image or binary format (TYPE I), which can be done by removing the particular extension (.txt in this example) from the EFT ASCII file list in the Copy/Move Action wizard in the **Advanced Options** dialog box. Refer to "Copy/Move (Push) File to Host Action" in the help for your version of EFT for the procedure.

GlobalSCAPE Knowledge Base
https://kb.globalscape.com/Knowledgebase/11113/Who-do-contents-of-a-file-ge...