

SOME CONCEPTS RELATED TO PROJECT SURFSTORE

George Porter
Feb 14, 2023

FILE METADATA (TOP) VS FILE DATA (BOTTOM)

```
gmporter@navygrog bin % stat -x protoc-gen-go
File: "protoc-gen-go"
Size: 8425842      FileType: Regular File
Mode: (0755/-rwxr-xr-x)  Uid: ( 501/gmporter)  Gid: ( 20/  staff)
Device: 1,15      Inode: 7741446    Links: 1
Access: Thu Jan 27 13:29:54 2022
Modify: Thu Jan 27 13:27:02 2022
Change: Thu Jan 27 13:27:02 2022
gmporter@navygrog bin %
```

```
gmporter@navygrog bin % hexyl protoc-gen-go | head -10
```

00000000	cf fa ed fe 0c 00 00 01	00 00 00 00 02 00 00 00	xxxx_00	00000000
00000010	0e 00 00 00 70 09 00 00	04 00 20 00 00 00 00 00	0000p_00	00 000000
00000020	19 00 00 00 48 00 00 00	5f 5f 50 41 47 45 5a 45	0000H000	__PAGEZE
00000030	52 4f 00 00 00 00 00 00	00 00 00 00 00 00 00 00	RO000000	00000000
00000040	00 00 00 00 01 00 00 00	00 00 00 00 00 00 00 00	00000000	00000000
00000050	00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	00000000	00000000
00000060	00 00 00 00 00 00 00 00	19 00 00 00 38 01 00 00	00000000	00008000
00000070	5f 5f 54 45 58 54 00 00	00 00 00 00 00 00 00 00	__TEXT00	00000000
00000080	00 00 00 00 01 00 00 00	00 80 31 00 00 00 00 00	00000000	0x100000

```
gmporter@navygrog bin %
```

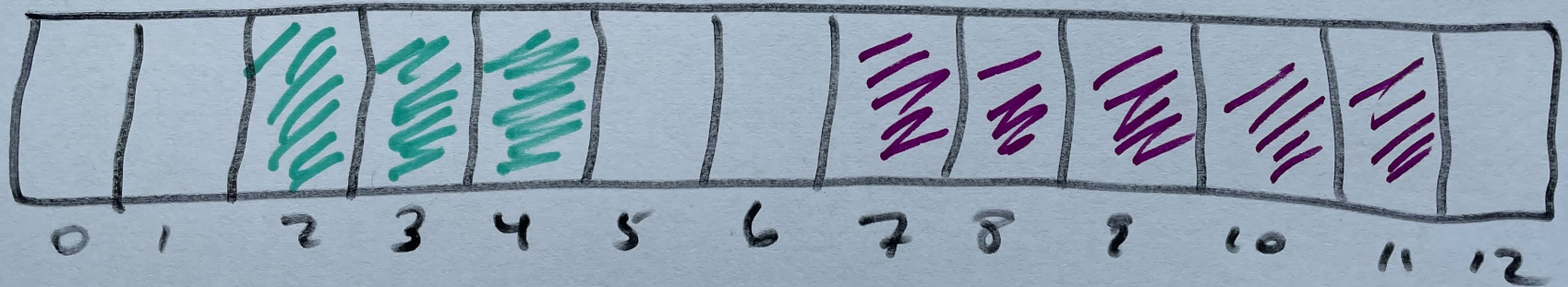
STORING FILE CONTENTS IN BLOCKS

File 1 (2, 3, 4) 12 KB

File 2 (7, 8, 9, 10, 11) 20 KB

File 3 (???) 16 KB

4 KB block



SEPARATING DATA FROM METADATA

- Metadata typically “points” to where the data is located
 - For local file systems *inodes* hold a list of *block addresses* (called LBA or logical block address)
- For network file systems like Dropbox, HDFS, Google File System (GFS) we need:
 - The IP address/hostname of the server that holds the block
 - A unique identifier for that block (often a hash value)

CRYPTOGRAPHIC HASH FUNCTIONS

- A hash function that maps from its input domain to its output domain in such a way that small changes to the input produce large changes in the output*

```
[gmporter@navygrog ~ % echo "hello CSE124" | shasum -a 256
36afb6d691a3d11af50398c5bd81eb080d9a039acca7911d6661712807eca72ea -
[gmporter@navygrog ~ % echo "Hello CSE124" | shasum -a 256
c02485236023d447856040c101df785285481992933d345dbccca7e6bbced258 -
gmporter@navygrog ~ % █
```

* It's more complicated than that.

THE SOURCE CODE TO THE UNITED STATES: THE US CONSTITUTION



IDENTIFYING BLOCKS OF DATA BY HASH FUNCTIONS

EXAMPLE: THE US CONSTITUTION

```
gmporter@navygrog surfstore % hexyl const.txt | head -20
```

00000000	57 65 20 74 68 65 20 50	65 6f 70 6c 65 20 6f 66	We the P	eo	ple of
00000010	20 74 68 65 20 55 6e 69	74 65 64 20 53 74 61 74	the Uni	ted	Stat
00000020	65 73 2c 20 69 6e 20 4f	72 64 65 72 20 74 6f 20	es, in O	rd	er to
00000030	66 6f 72 6d 20 61 20 6d	6f 72 65 20 70 65 72 66	form a m	ore	perf
00000040	65 63 74 20 55 6e 69 6f	6e 2c 0a 65 73 74 61 62	ect Unio	n,	estab
00000050	6c 69 73 68 20 4a 75 73	74 69 63 65 2c 20 69 6e	lish Jus	tice,	in
00000060	73 75 72 65 20 64 6f 6d	65 73 74 69 63 20 54 72	sure dom	estic	Tr
00000070	61 6e 71 75 69 6c 69 74	79 2c 20 70 72 6f 76 69	anquilit	y, provi	
00000080	64 65 20 66 6f 72 20 74	68 65 20 63 6f 6d 6d 6f	de for t	he commo	
00000090	6e 0a 64 65 66 65 6e 63	65 2c 20 70 72 6f 6d 6f	n_defenc	e, promo	
000000a0	74 65 20 74 68 65 20 67	65 6e 65 72 61 6c 20 57	te the g	eneral	W
000000b0	65 6c 66 61 72 65 2c 20	61 6e 64 20 73 65 63 75	elfare,	and secu	
000000c0	72 65 20 74 68 65 20 42	6c 65 73 73 69 6e 67 73	re the B	lessings	
000000d0	20 6f 66 20 4c 69 62 65	72 74 79 20 74 6f 0a 6f	of Libe	rty to_o	
000000e0	75 72 73 65 6c 76 65 73	20 61 6e 64 20 6f 75 72	urselves	and our	
000000f0	20 50 6f 73 74 65 72 69	74 79 2c 20 64 6f 20 6f	Posteri	ty, do o	
00000100	72 64 61 69 6e 20 61 6e	64 20 65 73 74 61 62 6c	rdain an	d establ	
00000110	69 73 68 20 74 68 69 73	20 43 6f 6e 73 74 69 74	ish this	Constit	
00000120	75 74 69 6f 6e 20 66 6f	72 20 74 68 65 0a 55 6e	ution fo	r the_Un	

```
gmporter@navygrog surfstore %
```

BLOCK IDS DETERMINED BY SHA-256 HASHES

```
gmporter@navygrog surfstore % cat printhashes.sh
#!/zsh

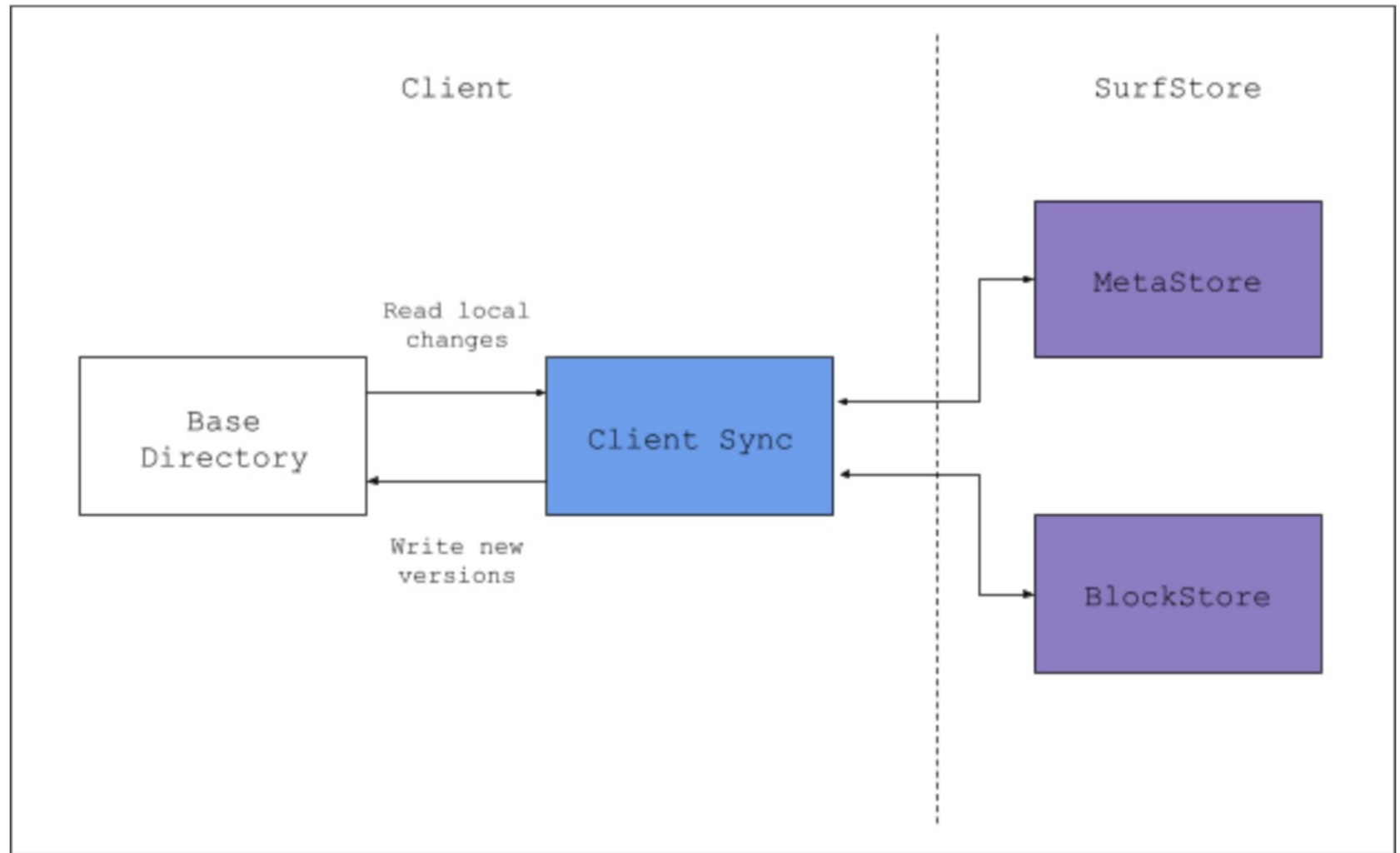
for i in `seq 0 5`;
do
    od -j $((8192 * $i)) -N 8192 const.txt | shasum -a 256
done
gmporter@navygrog surfstore %
```

```
gmporter@navygrog surfstore % ./printhashes.sh
672e9bfff6a0bc59669954be7b2c2726a74163455ca18664cc350030bc7eca71e -
31f28d5a995dcdb7c5358fcfa8b9c93f2b8e421fb4a268ca5dc01ca4619dfe5f -
172baa036a7e9f8321cb23a1144787ba1a0727b40cb6283dbb5cba20b84efe50 -
745378a914d7bcd26d3229f98fc2c6887e7d882f42d8491530dfaf4effef827 -
912b9d7afecb114fdaefecfa24572d052dde4e1ad2360920ebfe55ebf2e1818e -
3e6e0ff9279538e95d4818eb7f13846b26a984671f75ca668098906fbded2484 -
gmporter@navygrog surfstore %
```

What happens when we add a new Amendment to the constitution?

THE START OF A NETWORKED FILE SYSTEM...

- Metadata service
 - Maps files to size, list of block *hashes*
- Blockstore service
 - Maps a hash value to the contents of the block



For Dropbox, Google Drive, Microsoft Onedrive, etc, you can imagine there are many thousand BlockStore instances and a small number of MetaStores

WHAT DOES OUR CONSTITUTION LOOK LIKE?

Filename: const.txt

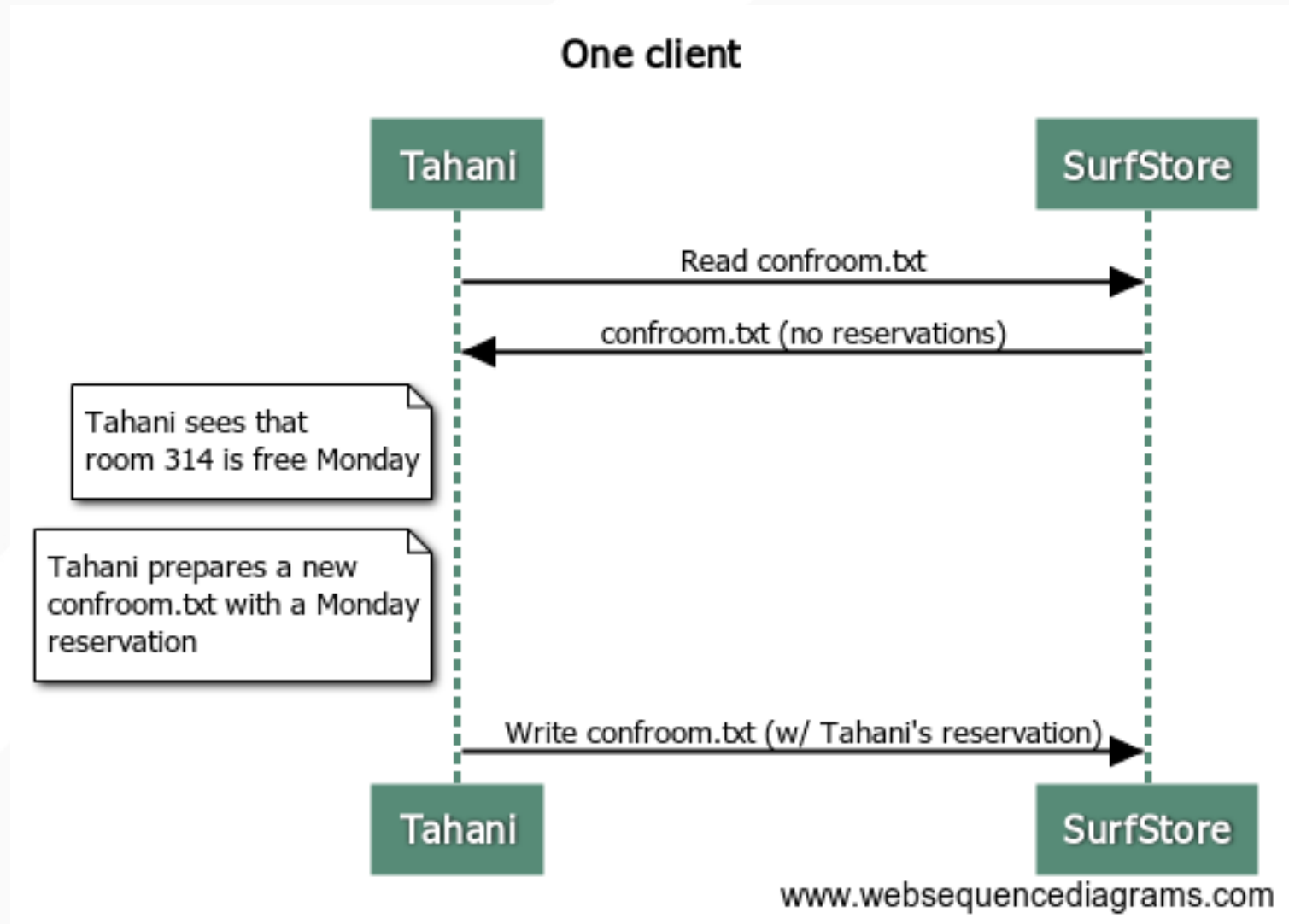
Size: 44841 bytes

Block hashes:

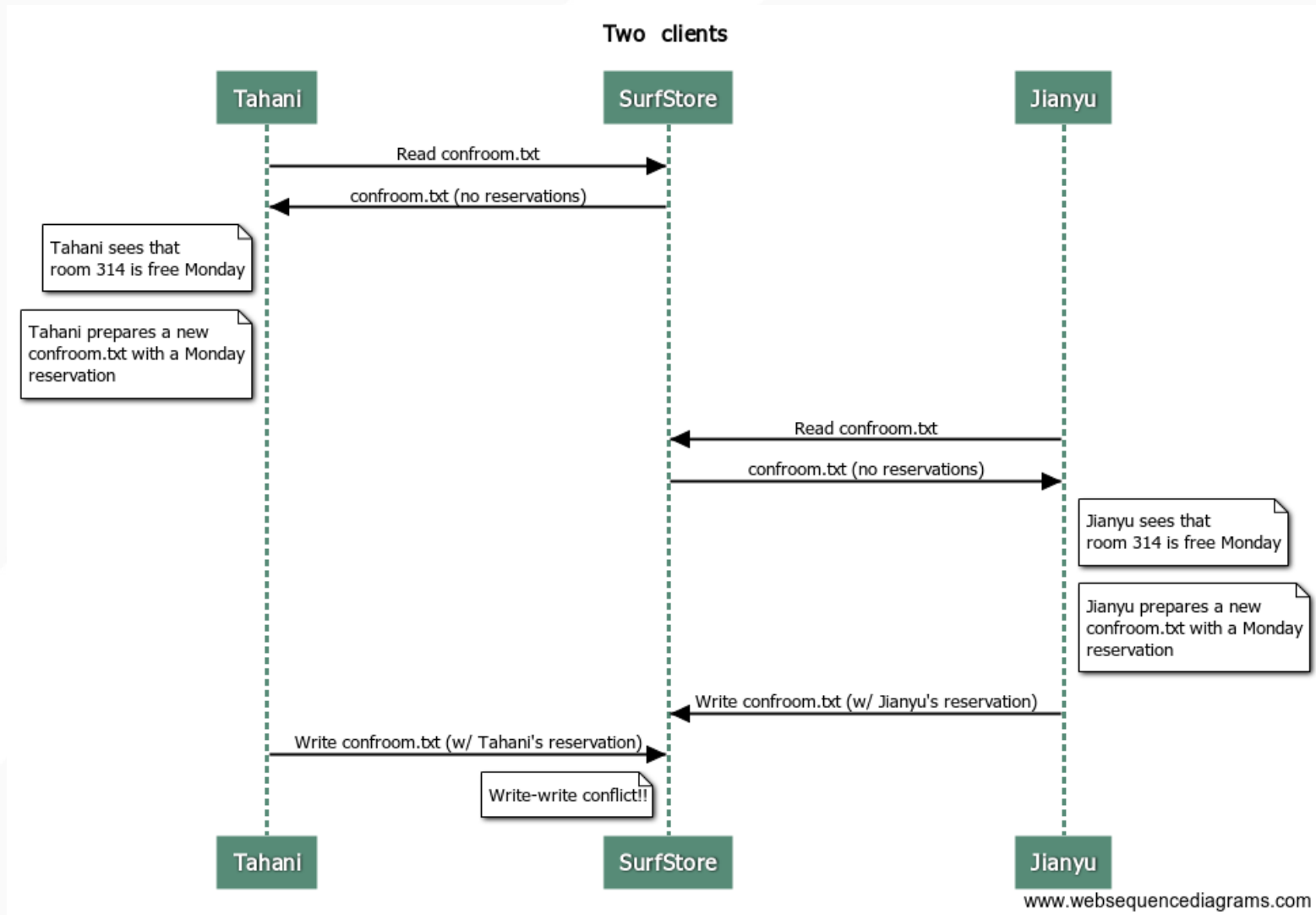
1. 672e9bffa0bc59669954be7b2c2726a74163455ca18664cc350030bc7eca71e
2. 31f28d5a995dcdb7c5358fcfa8b9c93f2b8e421fb4a268ca5dc01ca4619dfe5f
3. 172baa036a7e9f8321cb23a1144787ba1a0727b40cb6283dbb5cba20b84efe50
4. 745378a914d7bcd26d3229f98fc2c6887e7d882f42d8491530dfaf4effef827
5. 912b9d7afecb114fdaefecfa24572d052dde4e1ad2360920ebfe55ebf2e1818e
6. 3e6e0ff9279538e95d4818eb7f13846b26a984671f75ca668098906fbded2484

To modify a file, you add/remove/change one or more of the hash values and update the file size (after uploading any new blocks to the blockstore)

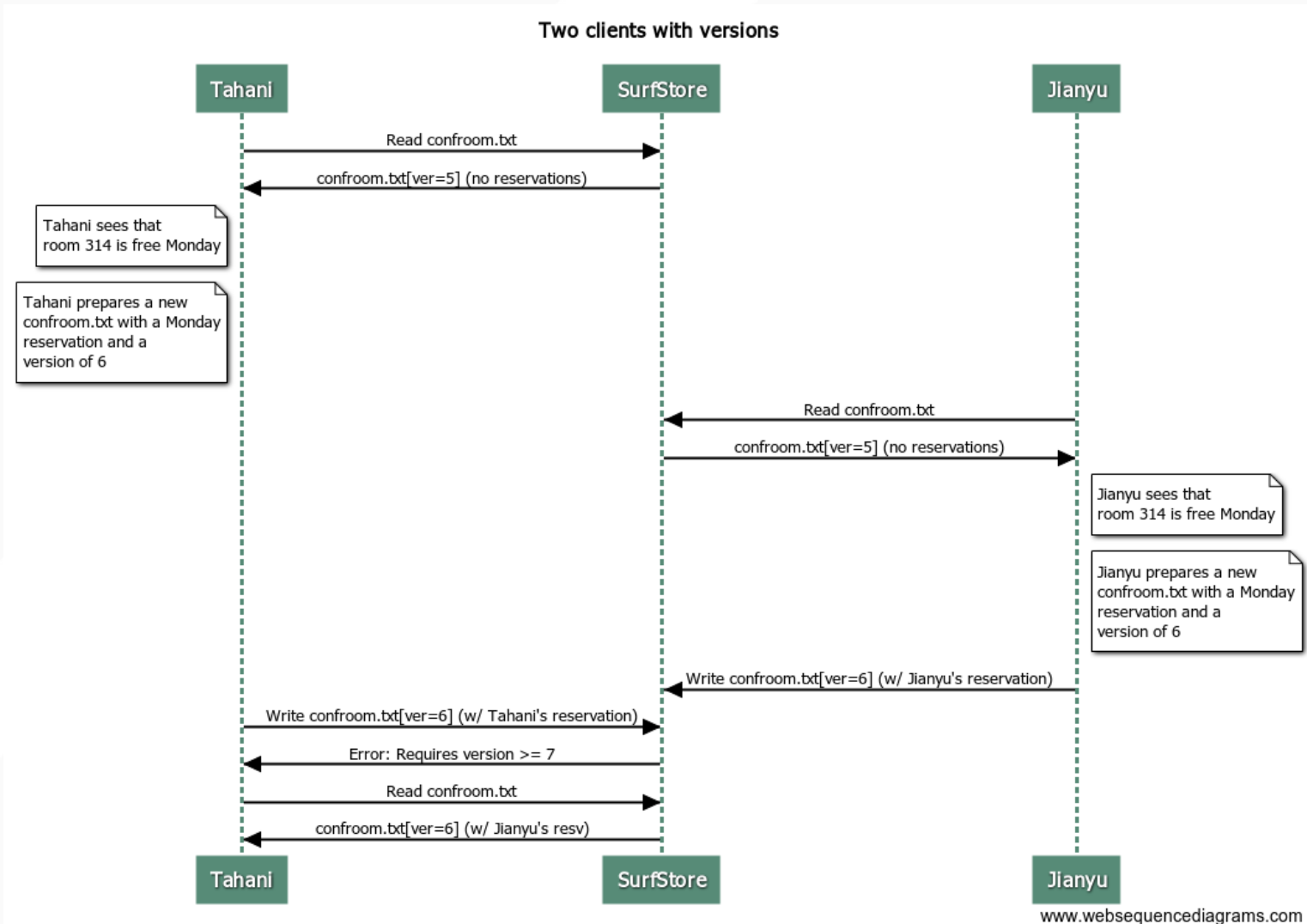
BUT A NETWORK SUPPORTS MULTIPLE USERS...



MULTIPLE USERS EDITING THE SAME FILE?



FILE VERSIONING



UC San Diego