

~~draft-prz-lsr-hierarchical-snps~~

draft-prz-lsr-ash-packets-00

Optimizing LSDB Sync at Scale

Tony P

Tony L

Joel H

HPE Juniper Networking



What's New?

- Well, a `_LOT_` since practical implementation mercilessly exposes holes in mental gymnastics-based specs
 - Joel joined and helped a lot to bend the spec into better shape as well so he has been cashiered as co-author
 - Renaming inspired by D. Bowie's Major Tom & Claude and then, who can dislike CASH
 - Turned out CSNP/PSNP model is needed
 - CASH is HASH equivalent of CSNP
 - PASH is equivalent to PSNP
 - J Snijders nudged us very nicely from Fletcher to Sip Hash
 - Going to SIP-1-3 64 bits to completely remove the collision discussion
 - Example added & Massive Restructuring
 - Subtle Clarifications on 0 Hash Value Semantics
-

An Example, My Horse for an Example

- About 3K Fragments, 100 nodes, ~300 Frags Diff
- Initial Cash Exchange Wild Mix of Things
- Mismatches Generate “Smaller Range” PASH
 - Unless 1 Node Only -> CSNP/PSNP/Flood
- Some Ranges Match + Hash Matches
- Some Ranges Match + Hash Mismatch
- Same with Range Mismatches
- Different Packings Due to Fragment Distribution

Node1→Node2

```
Pkt 1 [1010.0000.00.00 - 1010.0063.00.00] ( 35 hashes):
1010.0000.00.00 - 1010.0001.FF.FF: 89 frags, hash=$7A4FDCCA7E967402
1010.0002.00.00 - 1010.0004.FF.FF: 90 frags, hash=$0CAF4414E24B82C9
1010.0005.00.00 - 1010.0005.FF.FF: 113 frags, hash=$8186E01BB9AD1BBC
1010.0006.00.00 - 1010.0007.FF.FF: 64 frags, hash=$73B0979A0CC85195
1010.0008.00.00 - 1010.000B.FF.FF: 84 frags, hash=$8972B64DE9B4DA4F
1010.000C.00.00 - 1010.0010.FF.FF: 62 frags, hash=$316708E48767BCB3
1010.0011.00.00 - 1010.0011.FF.FF: 38 frags, hash=$44B8E43854718EBD PEER MISMATCH
1010.0012.00.00 - 1010.0012.FF.FF: 203 frags, hash=$E9514D9AF7A3E066
1010.0013.00.00 - 1010.0019.FF.FF: 92 frags, hash=$432EFCB7545F60C9 PEER MISMATCH
1010.001A.00.00 - 1010.001C.FF.FF: 43 frags, hash=$B6F8465659A1CEF5 PEER MISMATCH
1010.001D.00.00 - 1010.001D.FF.FF: 107 frags, hash=$35B83DF10500FED3 PEER MISMATCH
1010.001E.00.00 - 1010.0021.FF.FF: 84 frags, hash=$B3809E4CCD1CF39C PEER MISMATCH
1010.0022.00.00 - 1010.0023.FF.FF: 30 frags, hash=$B4BA6BAD8EA0965C
1010.0024.00.00 - 1010.0025.FF.FF: 82 frags, hash=$420E372BC8BD7A39 PEER MISMATCH
1010.0026.00.00 - 1010.0028.FF.FF: 70 frags, hash=$CB7ECDC8D74ECD7C
1010.0029.00.00 - 1010.002A.FF.FF: 73 frags, hash=$146218E1332308DF
1010.002B.00.00 - 1010.002B.FF.FF: 61 frags, hash=$7125DD2220DB94FF
```

Single Node -> PASH

Node2→Node1

```
Pkt 1 [1010.0000.00.00 - 1010.0063.00.00] ( 32 hashes):
1010.0000.00.00 - 1010.0001.FF.FF: 89 frags, hash=$7A4FDCCA7E967402
1010.0002.00.00 - 1010.0004.FF.FF: 90 frags, hash=$0CAF4414E24B82C9
1010.0005.00.00 - 1010.0005.FF.FF: 113 frags, hash=$8186E01BB9AD1BBC
1010.0006.00.00 - 1010.0007.FF.FF: 64 frags, hash=$73B0979A0CC85195
1010.0008.00.00 - 1010.000B.FF.FF: 84 frags, hash=$8972B64DE9B4DA4F
1010.000C.00.00 - 1010.0011.FF.FF: 62 frags, hash=$316708E48767BCB3 PEER MISMATCH
1010.0012.00.00 - 1010.0012.FF.FF: 203 frags, hash=$E9514D9AF7A3E066
1010.0013.00.00 - 1010.0018.FF.FF: 92 frags, hash=$432EFCB7545F60C9
1010.0019.00.00 - 1010.001E.FF.FF: 71 frags, hash=$ED653FB228DBC390 PEER MISMATCH
1010.001F.00.00 - 1010.0023.FF.FF: 91 frags, hash=$5A8EF0CAEF8B0388
1010.0024.00.00 - 1010.0025.FF.FF: 82 frags, hash=$F52D1BF15F989B46 PEER MISMATCH
1010.0026.00.00 - 1010.0028.FF.FF: 70 frags, hash=$CB7ECDC8D74ECD7C
1010.0029.00.00 - 1010.002A.FF.FF: 73 frags, hash=$146218E1332308DF
1010.002B.00.00 - 1010.002B.FF.FF: 61 frags, hash=$7125DD2220DB94FF
```

Node 1 Finds 9 Mismatches, 2 with 0 fragments -> CSNPs, Other 4 Cut in Half -> 4 More Than

Remaining PASHs After Resulting CSNPs/PSNPs and Flooding

Node1→Node2 (round cuts) → 1 PASH packet(s)

Pkt 1 [1010.000C.00.00 - 1010.001E.FF.FF] (4 hashes):

1010.000C.00.00 - 1010.000F.00.00:	51 frags, hash=\$0EF48698E8152334	
1010.0010.00.00 - 1010.0011.FF.FF:	49 frags, hash=\$7B2B6A443B03113A	PEER MISMATCH
1010.0019.00.00 - 1010.001D.00.00:	150 frags, hash=\$83407BA75CA13026	PEER MISMATCH
1010.001E.00.00 - 1010.001E.FF.FF:	23 frags, hash=\$5DB4052BAC376648	PEER MISMATCH

Node 1 Answers with 651 Fragments 7 PSNPs

Further Cutting of Mismatches 2 More PSNPs

Node2→Node1 (round cuts) → 1 PASH packet(s) above

Pkt 1 [1010.0013.00.00 - 1010.004A.FF.FF] (4 hashes):

1010.0013.00.00 - 1010.0017.00.00:	67 frags, hash=\$D5713B83C9B813B2	
1010.0018.00.00 - 1010.0019.FF.FF:	53 frags, hash=\$9C46D732761D42E3	PEER MISMATCH
1010.0047.00.00 - 1010.0048.00.00:	69 frags, hash=\$12E8A50E9124F446	PEER MISMATCH
1010.0049.00.00 - 1010.004A.FF.FF:	38 frags, hash=\$F1BDC50367228544	PEER MISMATCH

Node 2 Answers with 508 Fragments 6 PSNPs

Further Cutting of Mismatches 3 More PSNPs

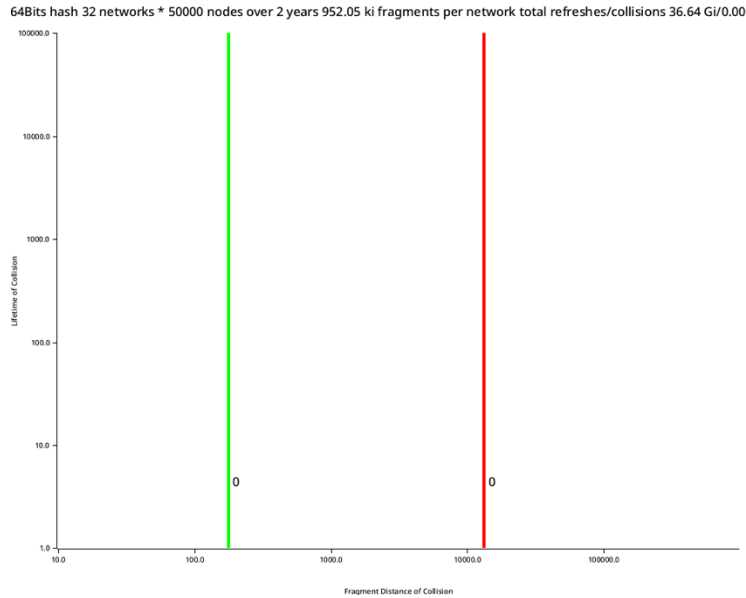
So Is It Worth It?

Metric	Node 1	Node 2	Total
ASH hash entries sent	39	36	75
PASH mismatch->resolved	4	4	8
~PASH PDUs (est.)	2	2	4
~PSNP PDUs (est., all frags)	9	9	18
Full-CSNP baseline PDUs	31	31	62
~Total PDUs (ASH+PSNP)	11	11	22
PDU ratio vs full CSNP	35.5%	35.5%	35.5%

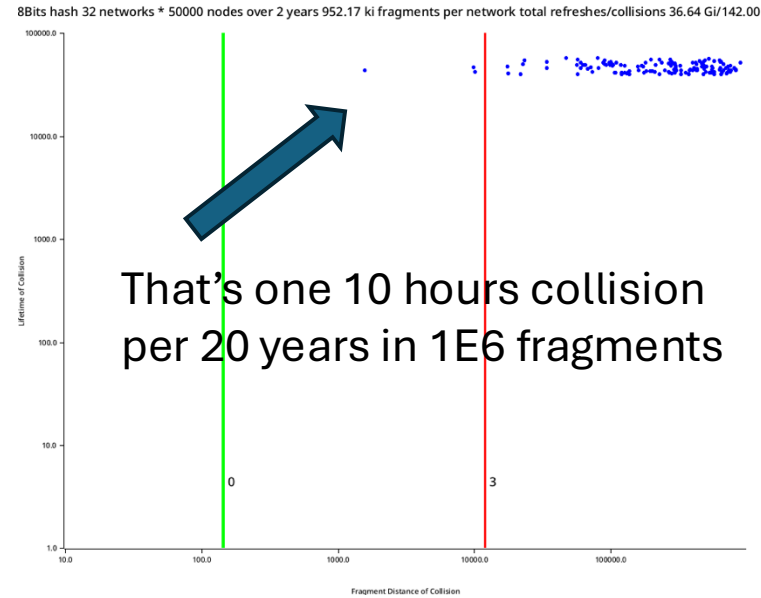
- Looks Decent For This Small Example
- ~85% Reduction For Something Like
 - 20K Fragments
 - 500 Nodes
 - 1500 Fragments Diff (7.5%)
- ~92% Reduction For Something Like
 - 100K Fragments
 - 15K Nodes
 - 500 Fragments Diff (0.5%)

So Why Flip to SIP-1-3 & 64-Bit Hashes

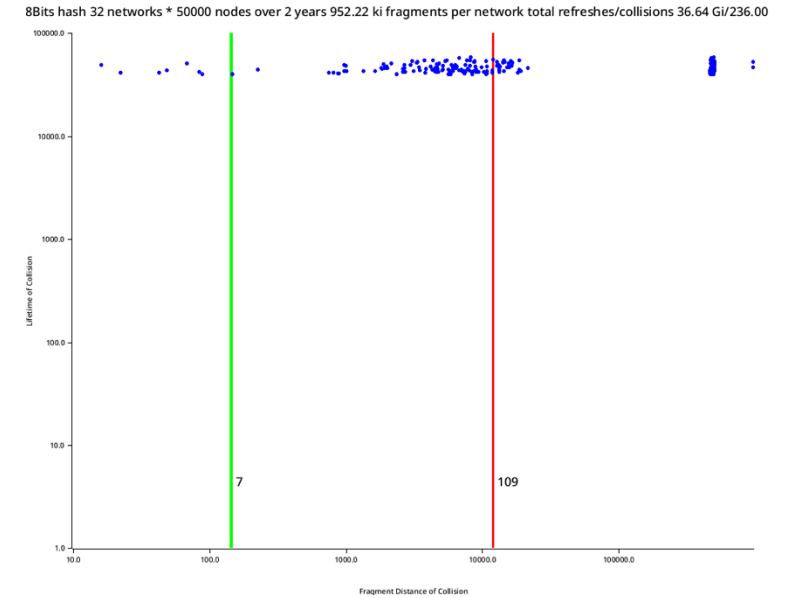
64-bits SIP



48-bits SIP



48-bits Fletcher



Red Line Single CASH, Green Line ~10 CASHes