



US 20210109682A1

(19) **United States**

(12) **Patent Application Publication**
Engel et al.

(10) **Pub. No.: US 2021/0109682 A1**

(43) **Pub. Date: Apr. 15, 2021**

(54) **EFFICIENT DATA COMPRESSION BY
GROUPING SIMILAR DATA WITHIN A
DATA SEGMENT**

(71) Applicant: **Pure Storage, Inc.**, Mountain View, CA
(US)

(72) Inventors: **Joern W. Engel**, Mountain View, CA
(US); **Yuhong Mao**, Fremont, CA (US)

(21) Appl. No.: **16/653,491**

(22) Filed: **Oct. 15, 2019**

Publication Classification

(51) **Int. Cl.**
G06F 3/06 (2006.01)
G06F 12/02 (2006.01)
H03M 7/30 (2006.01)

(52) **U.S. Cl.**

CPC **G06F 3/0661** (2013.01); **G06F 3/0608**
(2013.01); **H03M 7/30** (2013.01); **G06F**
3/0673 (2013.01); **G06F 12/0253** (2013.01);
G06F 3/064 (2013.01)

(57)

ABSTRACT

A plurality of data blocks stored at a plurality of solid-state storage devices are identified. A portion of data is read from each data block of the plurality of data blocks. A corresponding property is determined for each data block of the plurality of data blocks based on reading the portion of the data. A set of data blocks from the plurality of data blocks is identified, wherein each data block of the set of data blocks is associated with a first corresponding property. The set of data blocks is stored at a data segment.

