HDD,SSD等のデータ消去の課題とフォレンジック

第17回デジタル・フォレンジック・コミュニティ2020 in TOKYO
2020年12月7日（月）

アイフォレンセ日本データ復旧研究所（株）
AIFORENSE JAPAN DATA RECOVERY, INC.
下垣内 太（しもがいと だい）

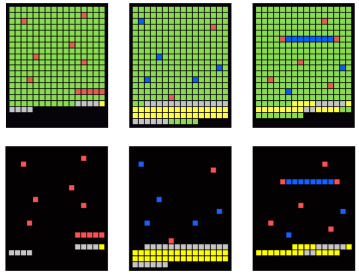
自己紹介

下垣内 太 （しもがいと だい）

大阪データ復旧 - 梅田・大阪駅前第4ビル
アイフォレンセ日本データ復旧研究所（株） 代表取締役
AIFORENSE JAPAN DATA RECOVERY, INC. (米国NY州、CEO)

一般社団法人 日本データ復旧協会 常任理事
デジタル・フォレンジック研究会 会員

愛知県豊川市出身
愛知県立時習館高等学校卒
関西大学総合情報学部卒
1998年創業



データ復旧 – 故障したコンピュータからデータを救出

デジタルフォレンジック調査 – 犯罪や不正事件の証拠品解析

https://www.facebook.com/dai.shimogaito

「 HDDの完全消去は不可能 」 朝日新聞2019年12月9日 夕刊

データ復元の達人 「リスク考え処分を」

消えたデータをよみがえらせる専門業者「大阪データ復旧」が大阪市にある。代表の下垣内太（しもがいと）さんは「HDDに記録されたデータを、100%安全に消去するのは不可能だ」と言い切る。

犯罪の容疑者が証拠隠滅のために消した記録装置からデータを取り出す「達人」として、下垣内さんには全国の警察から依頼が舞い込む。データを完全に消去する国際標準規格にのっとり初期化されたHDDから、消しきれなかったデータを復元させる手法を2016年、国際会議で発表し、話題になった。

そんな下垣内さんも「確実にデータを消す最も安全な方法は、HDDを破壊すること」と言う。

神奈川県のHDD流出問題について、「県庁内でディスクを破壊したり、職員立ち会いの下で外部で処理したりすれば、回避できた可能性がある。残ったデータが流出するリスクを避けるための運用基準を作り、処分方法を考えていけば、多くの企業や組織にありがちな『業者任せ』は起きないのではないか」と話す。

（編集委員・須藤龍也）

参照：朝日新聞デジタル
https://www.asahi.com/
Accessed 2019/12/07

２０１９年１２月６日　朝日新聞朝刊

元年）12月6日（金）　朝日新聞　1892年3月17日第3種郵便物認可　©朝日新聞社 2019年　49563号（日刊）

県秘密情報 大量流出

神奈川 税記録や名簿

HDD9個 廃棄業者社員が転売

納税などに関する大量の個人情報や秘密情報を含む神奈川県庁の行政文書が蓄積されたハードディスク（HDD）が、ネットオークションを通じて転売され、流出していたことが朝日新聞の取材で分かった。県のサーバーから取り外されたHDDのデータ消去が不十分なまま、中古品として出回っていた。県によると、データの消去から廃棄までを請け負った業者の社員が、転売に関与したことを認めているという。

転売されたHDDは縦約15㌢、横約10㌢、厚さ約2・5㌢。少なくとも9個あり、この中に保存されたデータの容量は27テラバイトに上る。仮に画像を添付したメール1通を3メガバイトとすると、900万通に相当する。神奈川県が調査を続けているが、情報流出の事案としては世界でもまれな規模に上る可能性がある。

県が確認したところ、HDDは県庁内の各部局の情報を蓄積する共有サーバーに使われていた。中には、法人名が記載された税務調査後の通知や、個人名や住所が記載された自動車税の納税記録、企業の提出書類、県職員の業務記録や名簿類などが含まれていた。

県によると、転売されていたHDDは、県が富士通リース（東京都千代田区）から借りたサーバーに使われたもので、今春に交換時期を迎え、サーバーから取り外された。富士通リースは県との契約に基づき、データを復元不可能な状態にする作業を、情報機器の再生事業を手がけるブロードリンク（同中央区）に委託。同社に対し富士通リースは、破壊して作動しないようにしてから廃棄するか、データを完全に消去するよう指示していた。

県からブロードリンクに引き渡された時点で、HDDには簡易なデータ消去（初期化）が施されていた。HDDは都内にあるブロードリンクの施設で保管されていたが、データの消去作業の担当者が一部を持ち出し、オークションサイトに出品したという。

出品されていたHDD9個を、IT企業経営の男性が仕事に使おうと落札。使用前に安全性を確かめるため男性が中身を確認したところ、データの存在に気づいた。復元ソフトを使うと、神奈川県の公文書とみられる大量のファイルが保存されていたという。

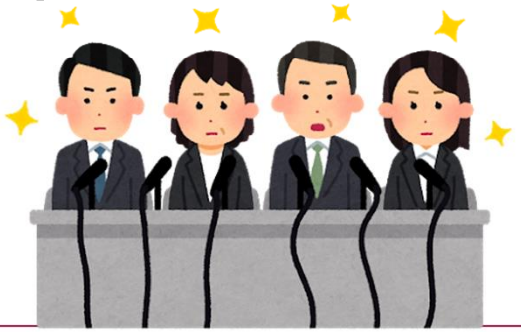
男性からの情報提供を受け、朝日新聞が11月27日に県に指摘。HDDに記されていた製品番号から、県のサーバーに使われていた実物と分かった。富士通リースは「現時点でコメントできることはない」。ブロードリンクの幹部は取材に対し、流出があったことを認

いた。復元ソフトを使うと、神奈川県の公文書とみられる大量のファイルが保存されていたという。

2019年（令和元年）
12月6日
金曜日
朝日新聞
朝日新聞大阪本社　〒530-8211 大阪市北区中之島2-3-18 電話 06-6231-0131 www.asahi.com

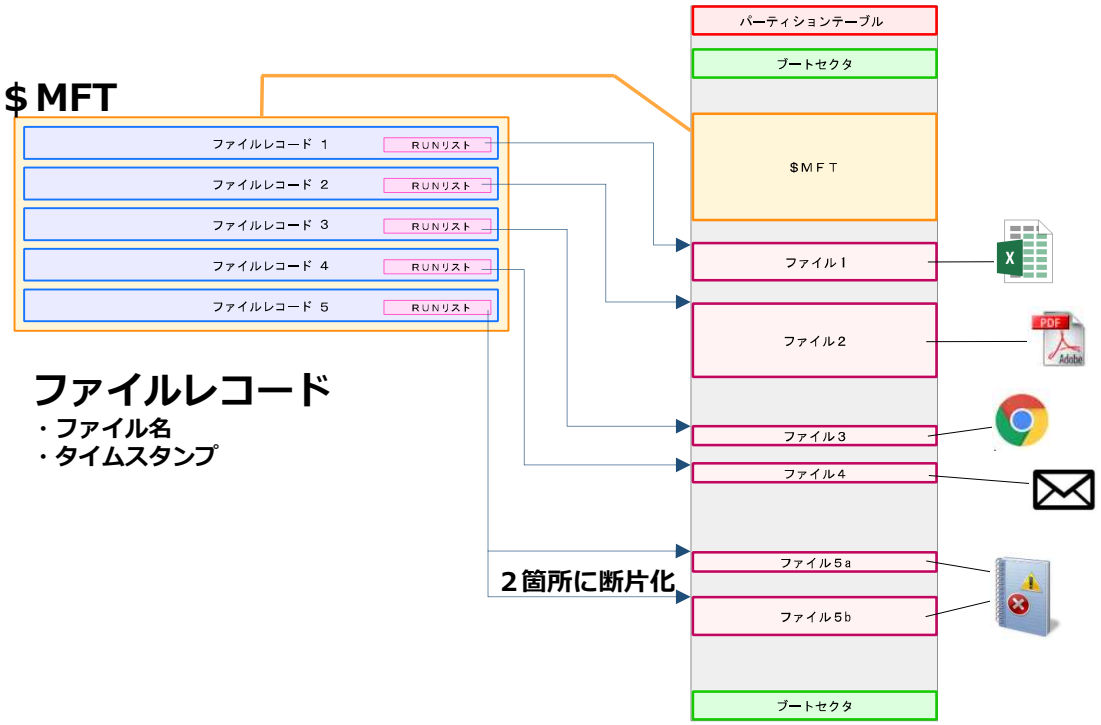
２０１９年１２月5日　神奈川県記者会見（コメント抜粋）

- （県）完全な形ではなかったり、壊れていたりというものもありましたので
- （県）ファイル名がすでに別のものになっておりますので
- （県）１２３４５６みたいなファイル名で開けてみるとなんかの統計データが出てきたり
- （県）オリジナルのものだと思われるようなタイトルのものは、無かったですよね
- （県）明らかに不自然なファイル名だなというのは分かりました
- （記者）壊れて開けないファイルは、割合的にはどれくらいなんですか？
  （県）結構あったよね。まず画像ファイルはほとんどだめでした。



※上記テキストは、以下のYoutubeサイトを閲覧し下垣内太が書き起こしたもの
【ノーカット】文書流出、神奈川県の黒岩祐治知事が会見「想定外だった」
https://www.youtube.com/watch?v=c4tV2q-md9Q　(Accessed : 2020/11/9)

ファイル復元の主な２つの解析方法の比較（可能性について）

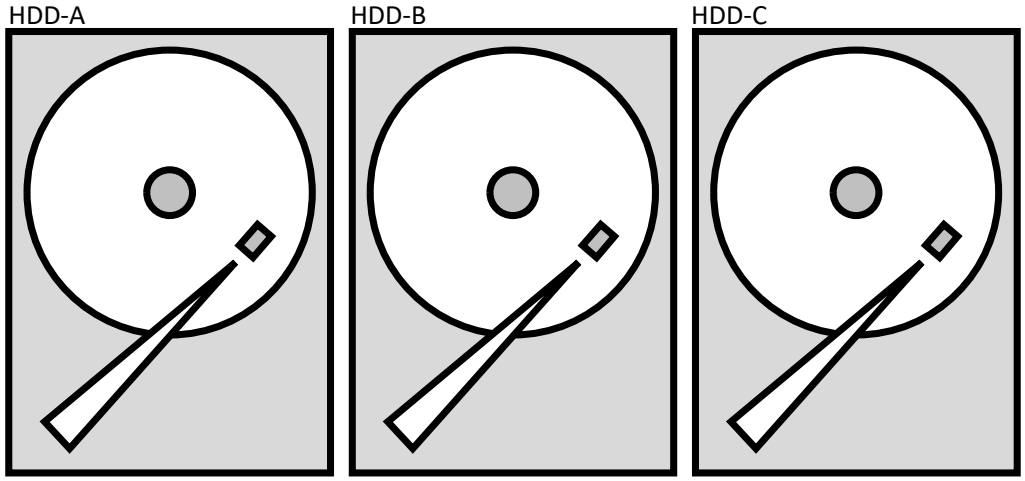


NTFSのデータ配置

	ファイルシステム解析	ファイルカービング解析
ファイル名の復元	○	×
ディレクトリ構造の復元	○	×
タイムスタンプの復元	○	×
ファイルの復元	• ファイルシステムの情報に基づいて出力される • ファイルの情報だけに依存した検出方法ではない	• ファイルの情報（主にファイルヘッダ）に基づいて検出（復元）される • ファイルシステムの情報には依存しない
特徴１	• ファイル名は復元されても、ファイルが正常復元するとは限らない。 • ディレクトリ構造、ファイル名、ファイル本体など、すべてを復元（再現）できることがある。	• ファイル名は復元されない • ファイル名は復元ソフトが付ける • 誤検出もある。むしろそれは前提 • ファイルシステム情報に不足があるときにも有効な解析方法 • ファイルシステム情報も解析対象になり得る
特徴２	• **流出情報の概要把握に有効**	• **流出情報の中身の把握に有効**

RAIDのデータ記録のしくみ

RAIDに記録されるファイルについて



HDD3台でRAID0構成（ストライプサイズ64KB）

JPG画像ファイルが3つ

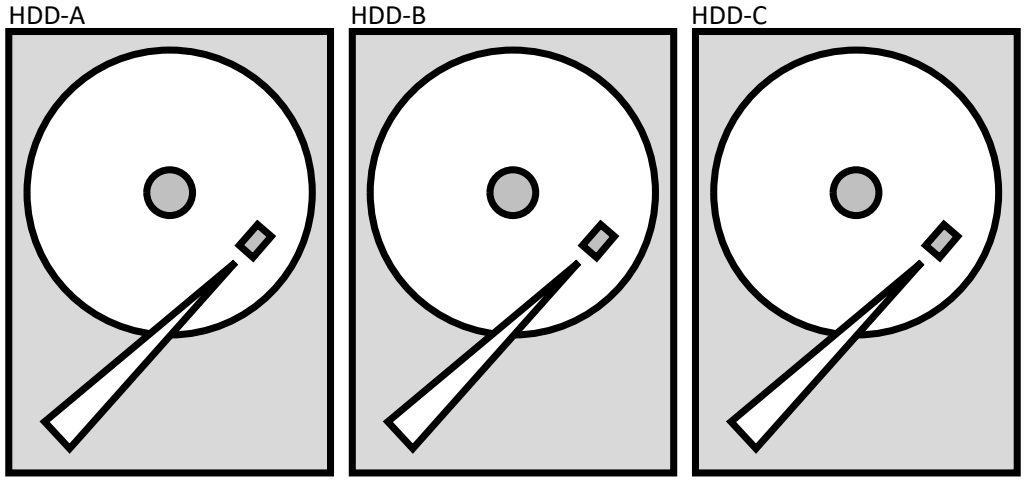
64KB

128KB

512KB

RAIDのデータ記録のしくみ

ファイルは、ストライプサイズずつ分割される

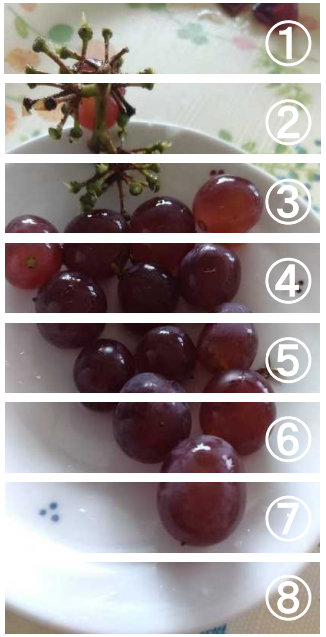


HDD3台でRAID0構成（ストライプサイズ64KB）

JPG画像ファイルが3つ

64KB

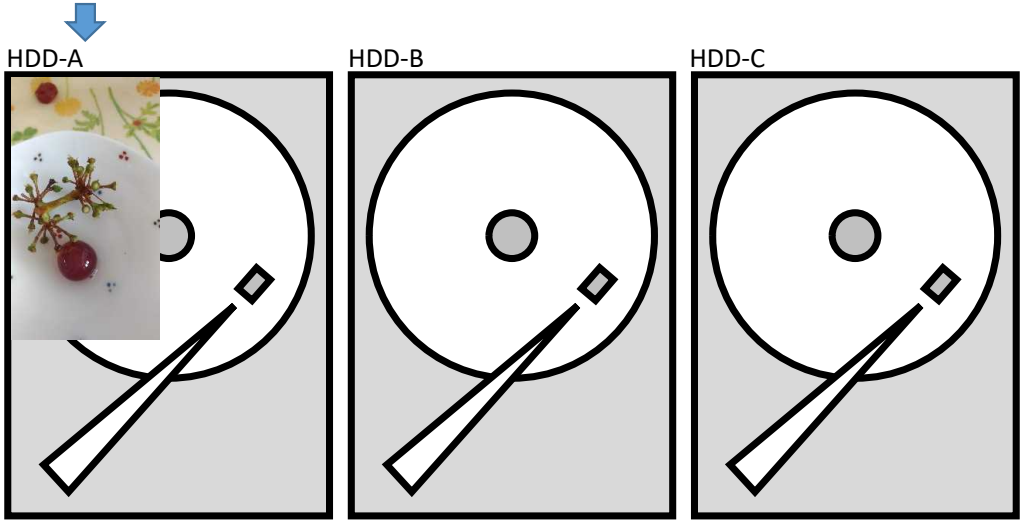
128KB
= 64KB x 2



512KB
= 64KB x 8

RAIDのデータ記録のしくみ

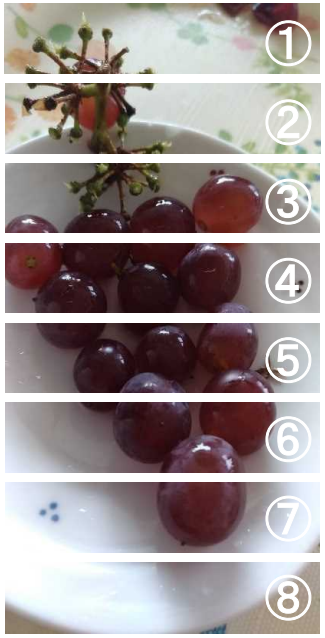
64KBのファイルは、分割されずに記録



HDD3台でRAID0構成（ストライプサイズ64KB）

64KB

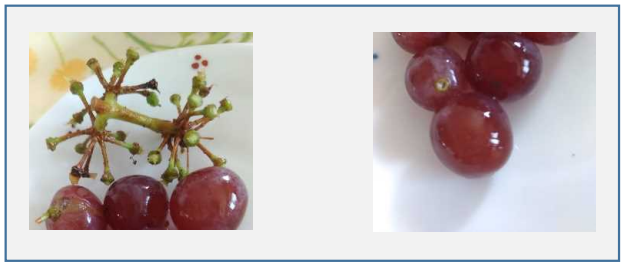
128KB
= 64KB x 2



512KB
= 64KB x 8

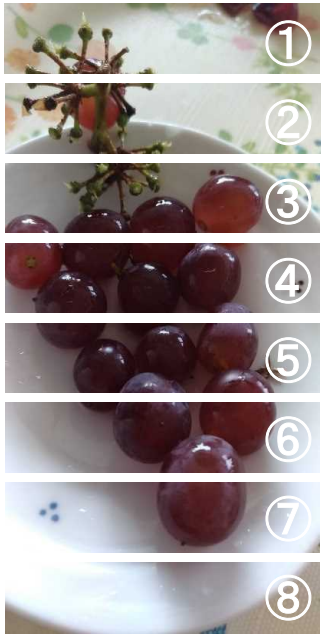
RAIDのデータ記録のしくみ

128KBのファイルは、2分割されて記録

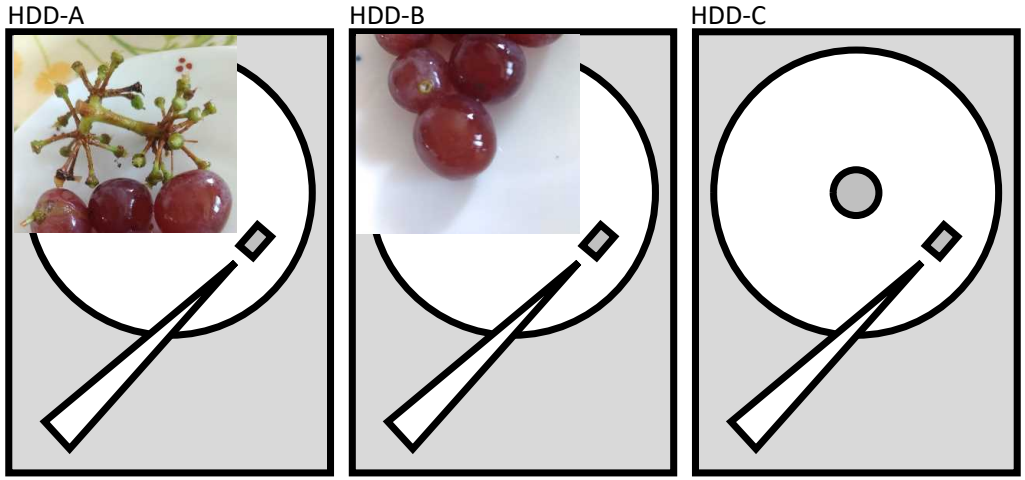


64KB

128KB
= 64KB x 2



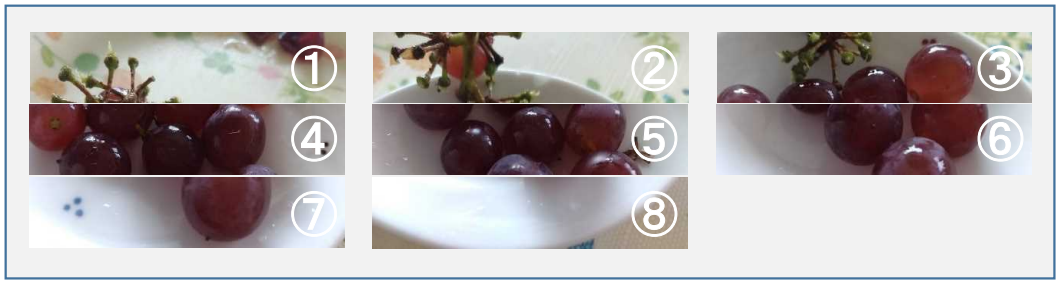
512KB
= 64KB x 8

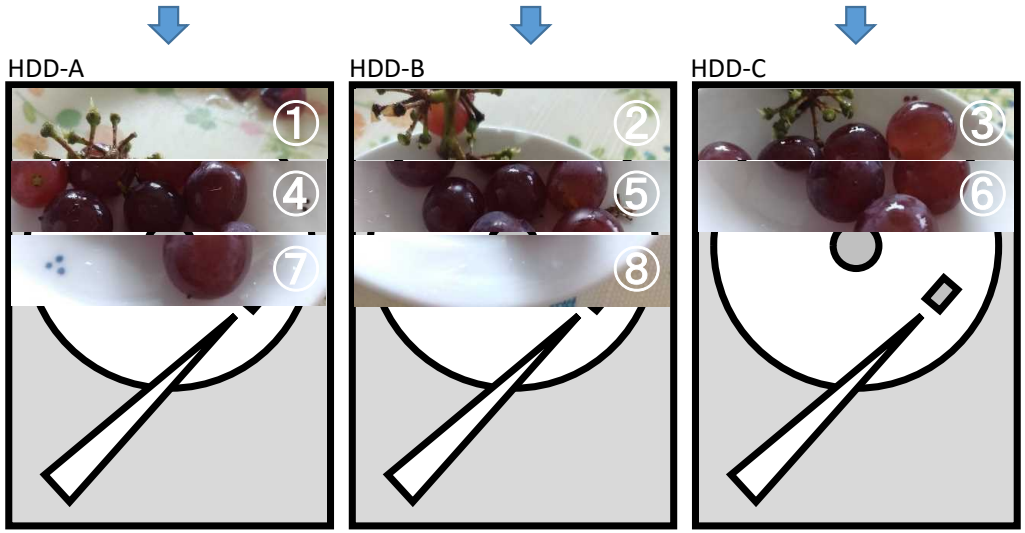


HDD3台でRAID0構成（ストライプサイズ64KB）

RAIDのデータ記録のしくみ

512KBのファイルは、8分割されて記録

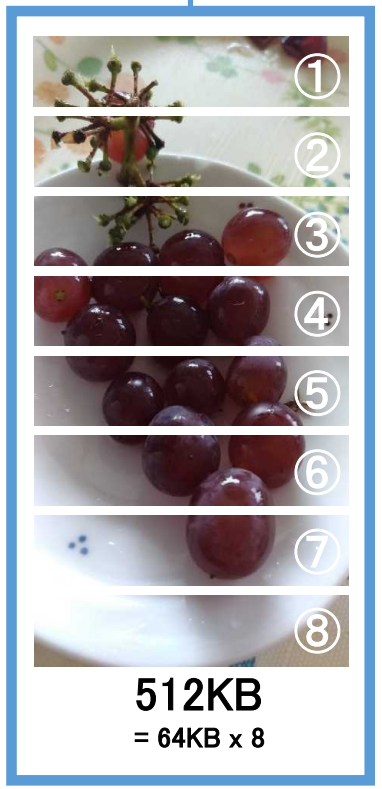




HDD3台でRAID0構成（ストライプサイズ64KB）

64KB

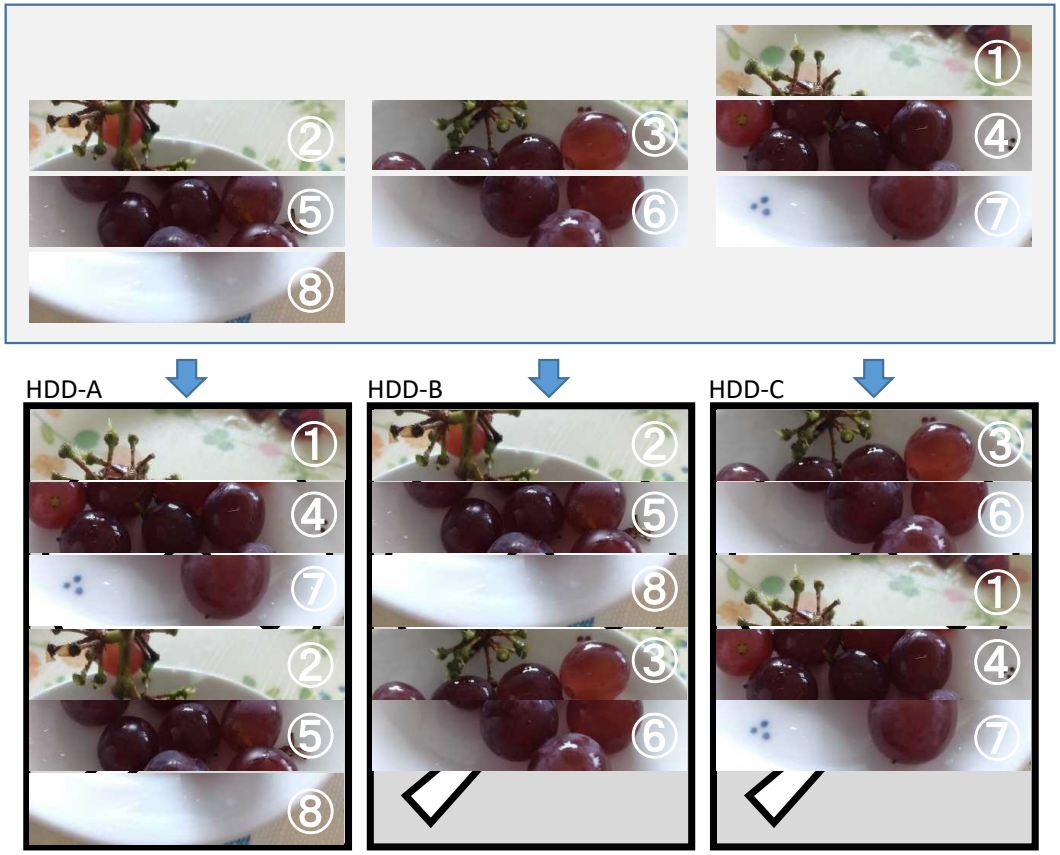
128KB
= 64KB x 2



512KB
= 64KB x 8

RAIDのデータ記録のしくみ

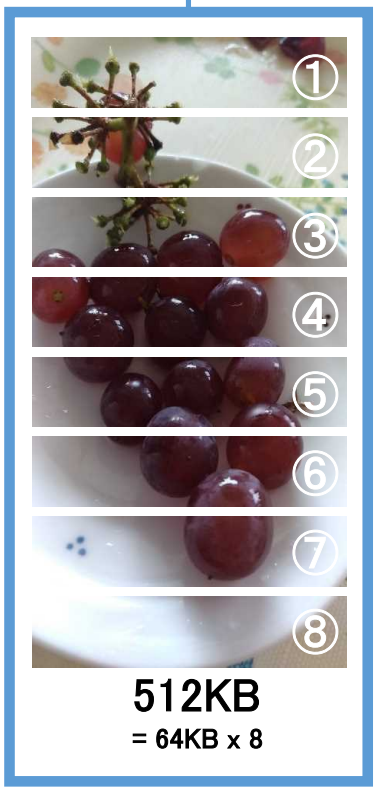
512KBの同一ファイルをさらに追加記録



HDD3台でRAID0構成（ストライプサイズ64KB）

64KB

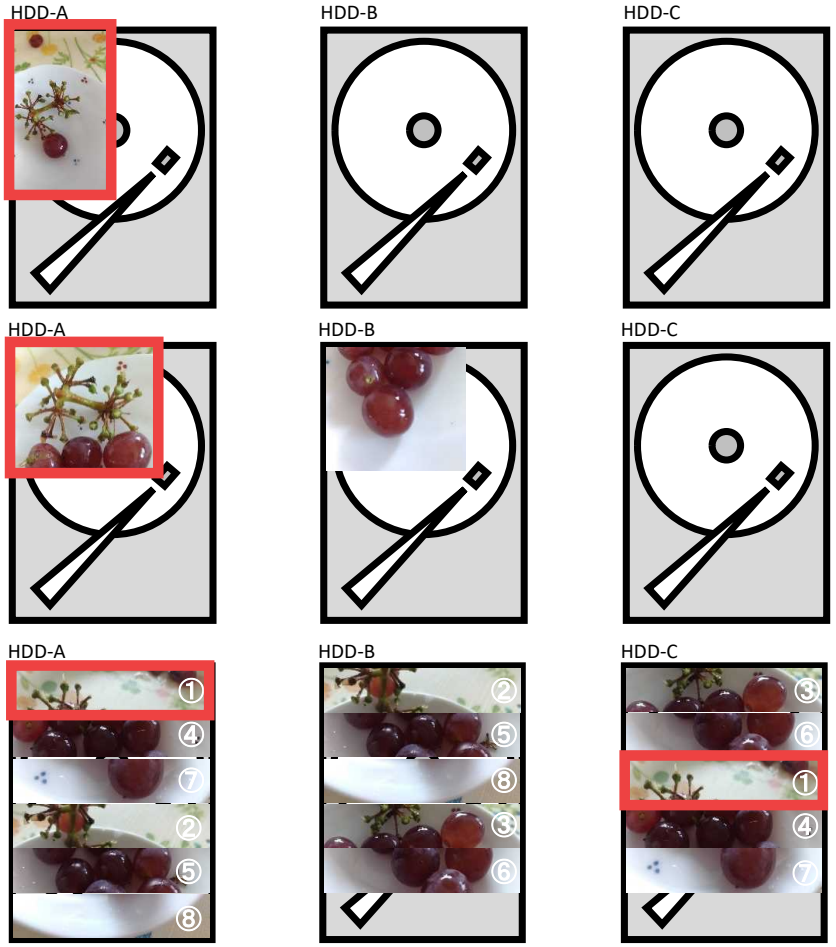
128KB
= 64KB x 2



512KB
= 64KB x 8

RAIDを構成するHDD単体をファイルカービング解析

ファイル容量とファイルヘッダ位置が復元結果を左右する



- ファイル容量がストライプサイズ以下

ファイルは正常に復元でき得る

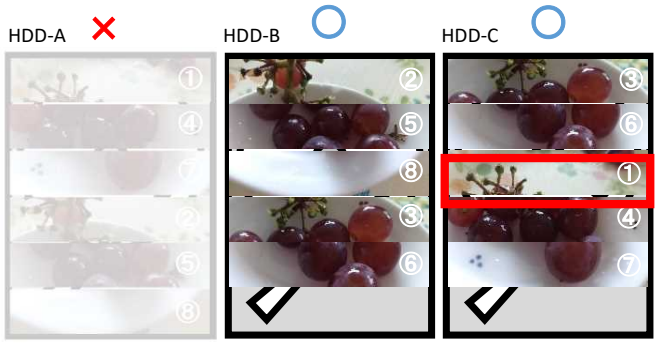
- ファイル容量がストライプサイズより大きい

ファイルヘッダを含む断片のみ検出される
よってファイルは不完全に復元される

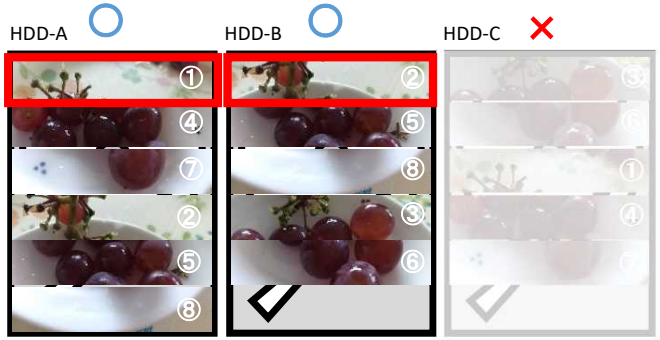
RAIDを構成するHDD複数台をファイルカービング解析

解析対象HDDの組み合わせにより復元データの数量が変わる

1ブロック
1ファイル
64KB分復元

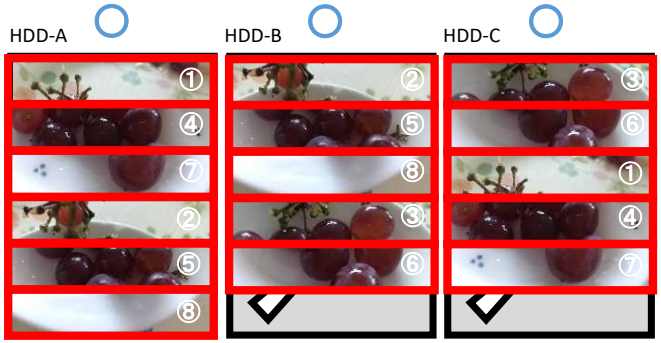


2ブロック
1ファイル
128KB分復元



3ブロック
2ファイル
192KB分復元

全ブロック
2ファイル正常復元

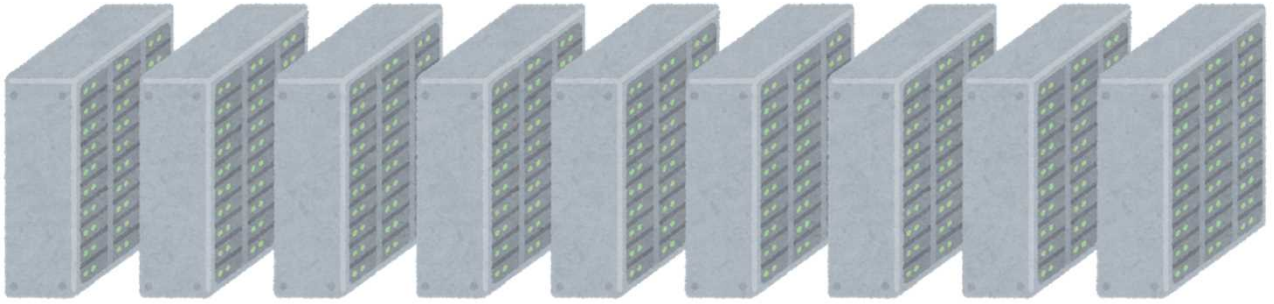


RAIDを構成する３６台のうち２９台を解析するイメージ図

RAIDの構成ドライブが全台なくても、復元できるデータはある

RAIDのデータは復元できる？

RAIDのデータは、復元できないんですよね？



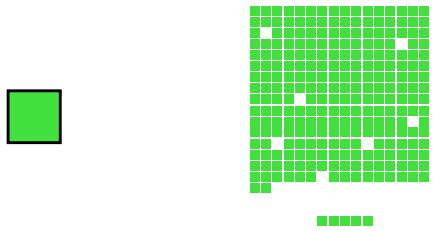
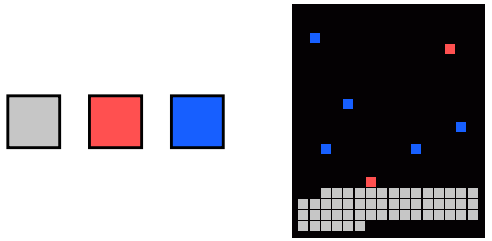
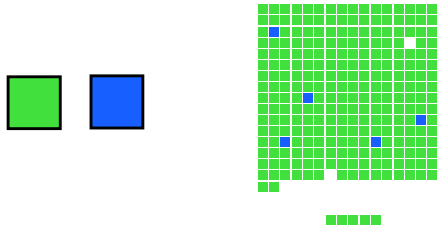
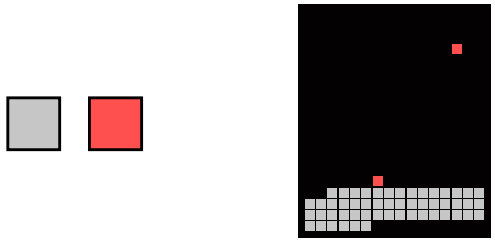
RAIDを構成するHDDに何らかの情報欠落がある場合のデータ復旧

ファイルシステム解析も実施すれば、ファイル名やタイムスタンプも復元

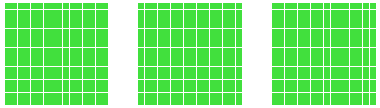
- **流出データ調査**の際は、**ファイル名、フォルダ、タイムスタンプ**などの概要情報の早期把握も**重要視**される
- ファイルカービング解析結果を、ファイルシステム解析に活用
- RAID0だけでなく、RAID5、RAID6、JBODでも解析可能
- HDDの並び順や、ストライプサイズは仕様書が無くても算出可能
- 欠損したファイルシステムデータは、仕様に沿って補填することで、より多くのファイル復元が期待できる

	ファイルシステム解析	ファイルカービング解析
ファイル名の復元	○	×
ディレクトリ構造の復元	○	×
タイムスタンプの復元	○	×
ファイルの復元	• ファイルシステムの情報に基づいて出力される • ファイルの情報だけに依存した検出方法ではない	• ファイルの情報（主にファイルヘッダ）に基づいて検出（復元）される • ファイルシステムの情報には依存しない
特徴1	• ファイル名は復元されても、ファイルが正常復元するとは限らない。 • ディレクトリ構造、ファイル名、ファイル本体など、すべてを復元（再現）できることがある。	• ファイル名は復元されない • ファイル名は復元ソフトが付ける • 誤検出もある。むしろそれが前提。 • ファイルシステム情報に不足があるときにも有効な解析方法 • ファイルシステム情報も解析対象になり得る
特徴2	• **流出情報の概要把握に有効**	• **流出情報の中身の把握に有効**

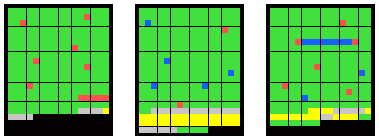
HDD - データ消去の対象と非対象領域

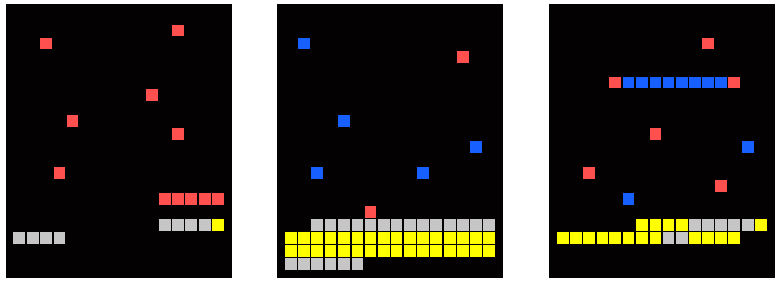
	消去可	消去不可
データ消去ソフト・ツール Secure Erase DoD方式（米国国防総省） グートマン方式（35回） ※物理破壊は除く		
Enhanced Secure Erase （NIST）		

悩ましい問題



HDDのデータ消去





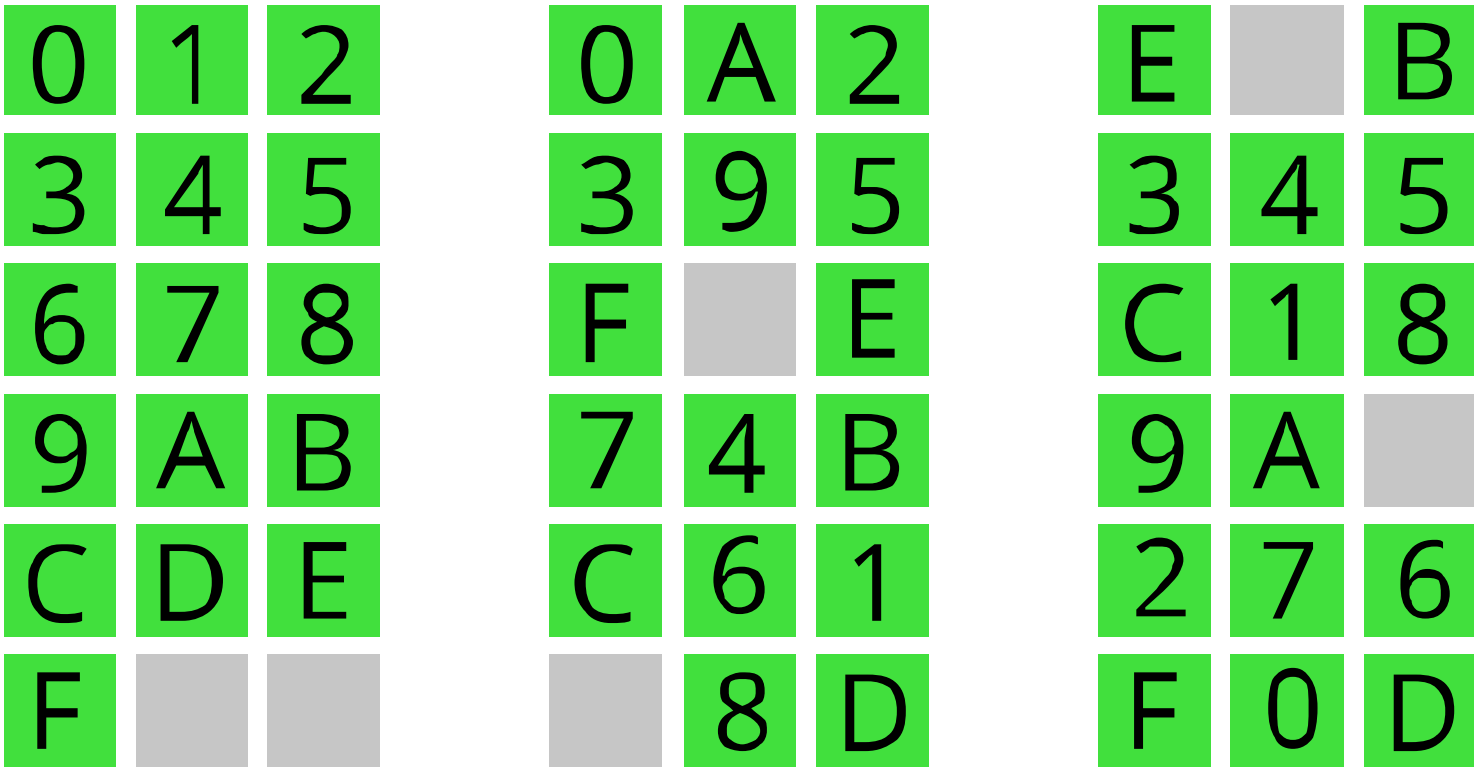
全セクタを**完全消去**できる

ソフトウェアは**無い**

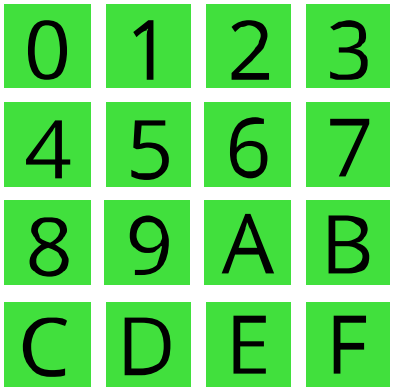
より詳細と実演映像は、以下URLをご覧ください

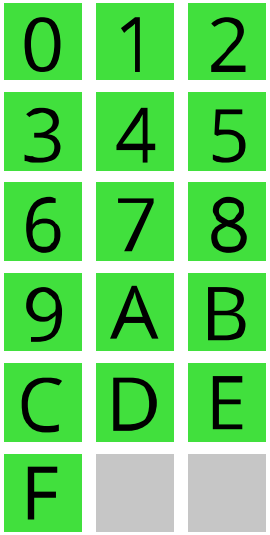
[CB16] EXOTIC DATA RECOVERY & PARADAIS by Dai Shimogaito
https://www.youtube.com/watch?v=llJkFJTFKLY

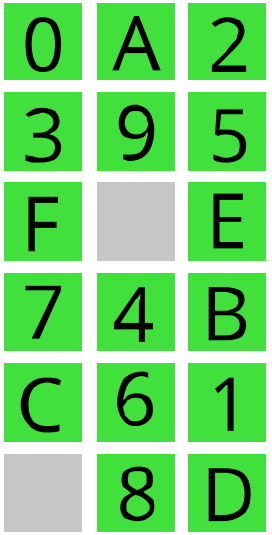
SSD：論理セクタの物理位置は流動的

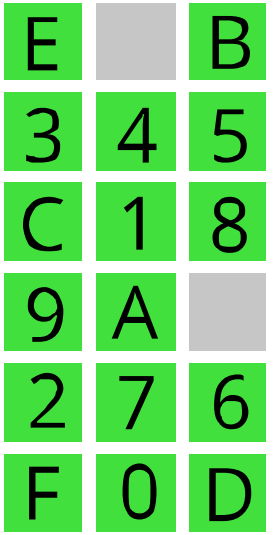


SSD ：論理位置と物理位置の比較

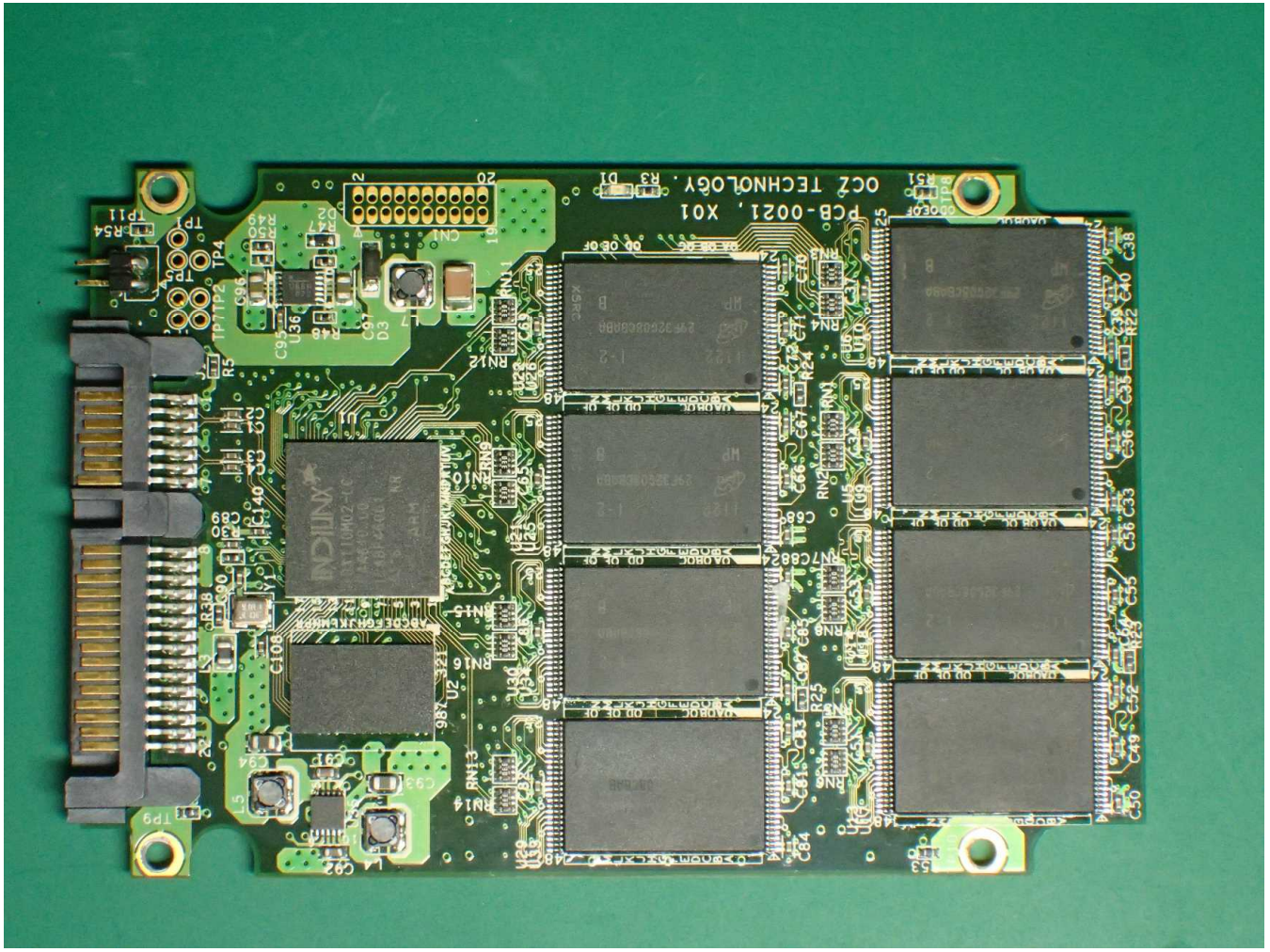




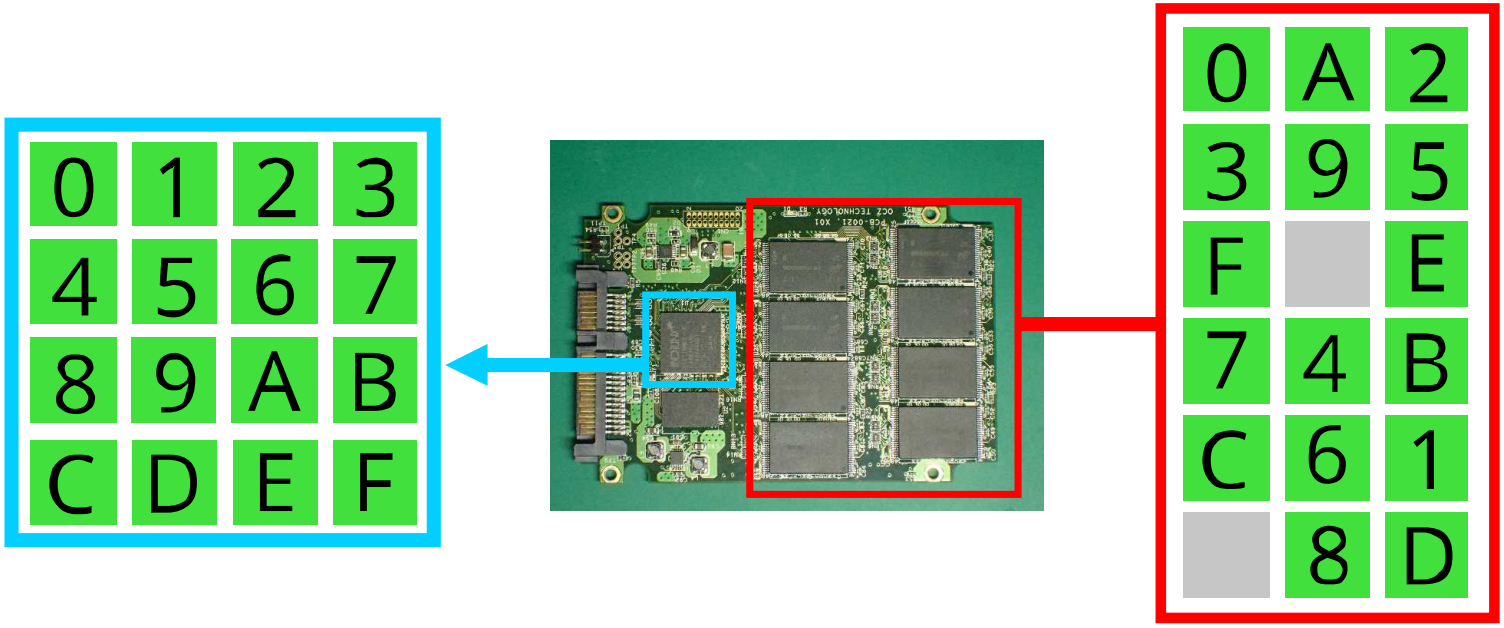




SSD基板のコントローラとNANDメモリチップ

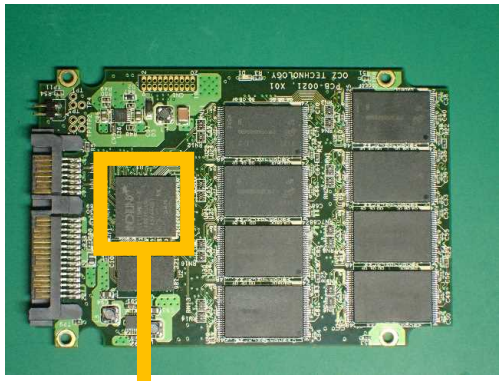


SSD ：論理位置と物理位置の比較



SSD :コントローラの役割

0	1	2	3
4	5	6	7
8	9	A	B
C	D	E	F



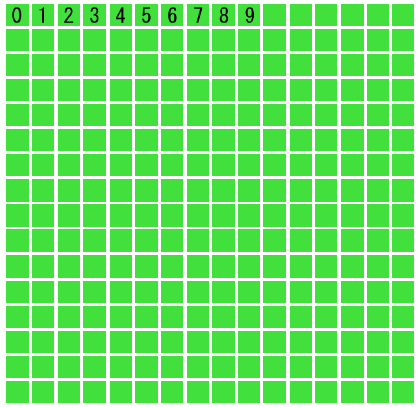
0	A	2
3	9	5
F		E
7	4	B
C	6	1
	8	D

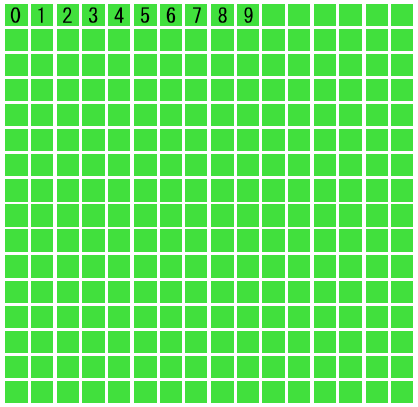
ファイルシステム

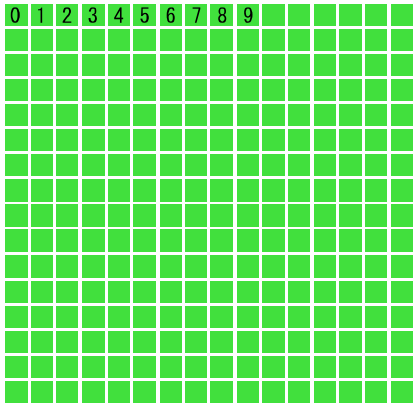
FTL （ Flash Translation Layer)
1) アドレス変換　※LBA <-> PBA
2) ウェアレベリング　※耐久性、速度
3) ガベージコレクション　※バックグラウンドでの消去準備
4) ハウスキーピング　※バックグラウンドでの消去実施

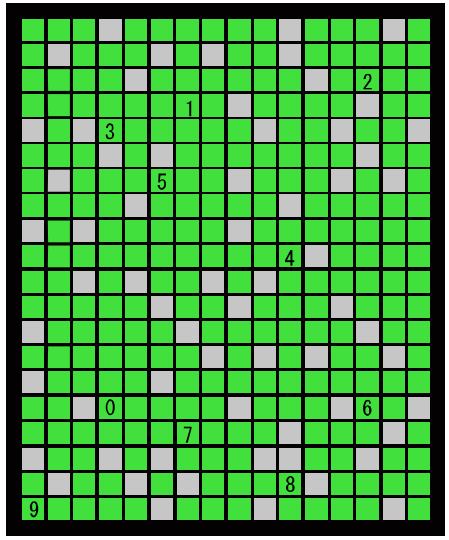
NAND Flash Memory

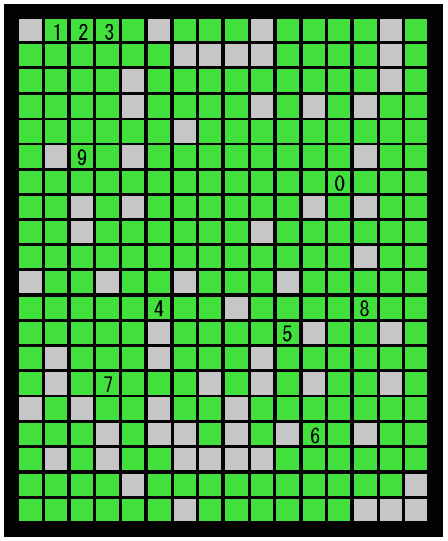
SSD ：論理セクタの物理位置は流動的

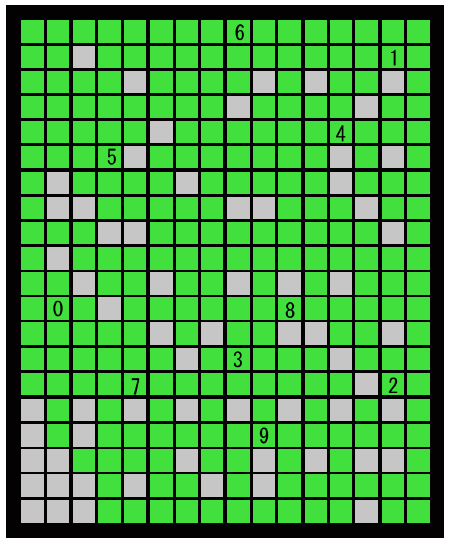




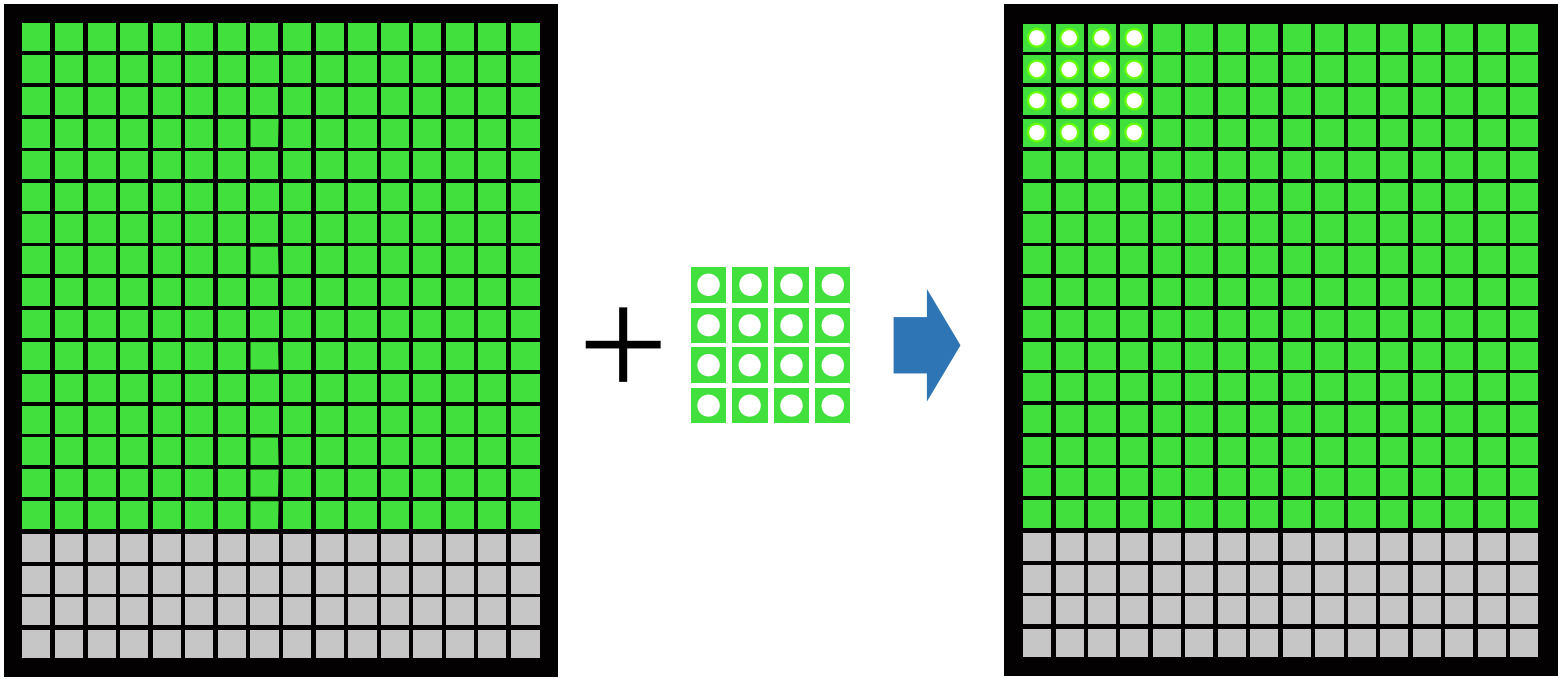




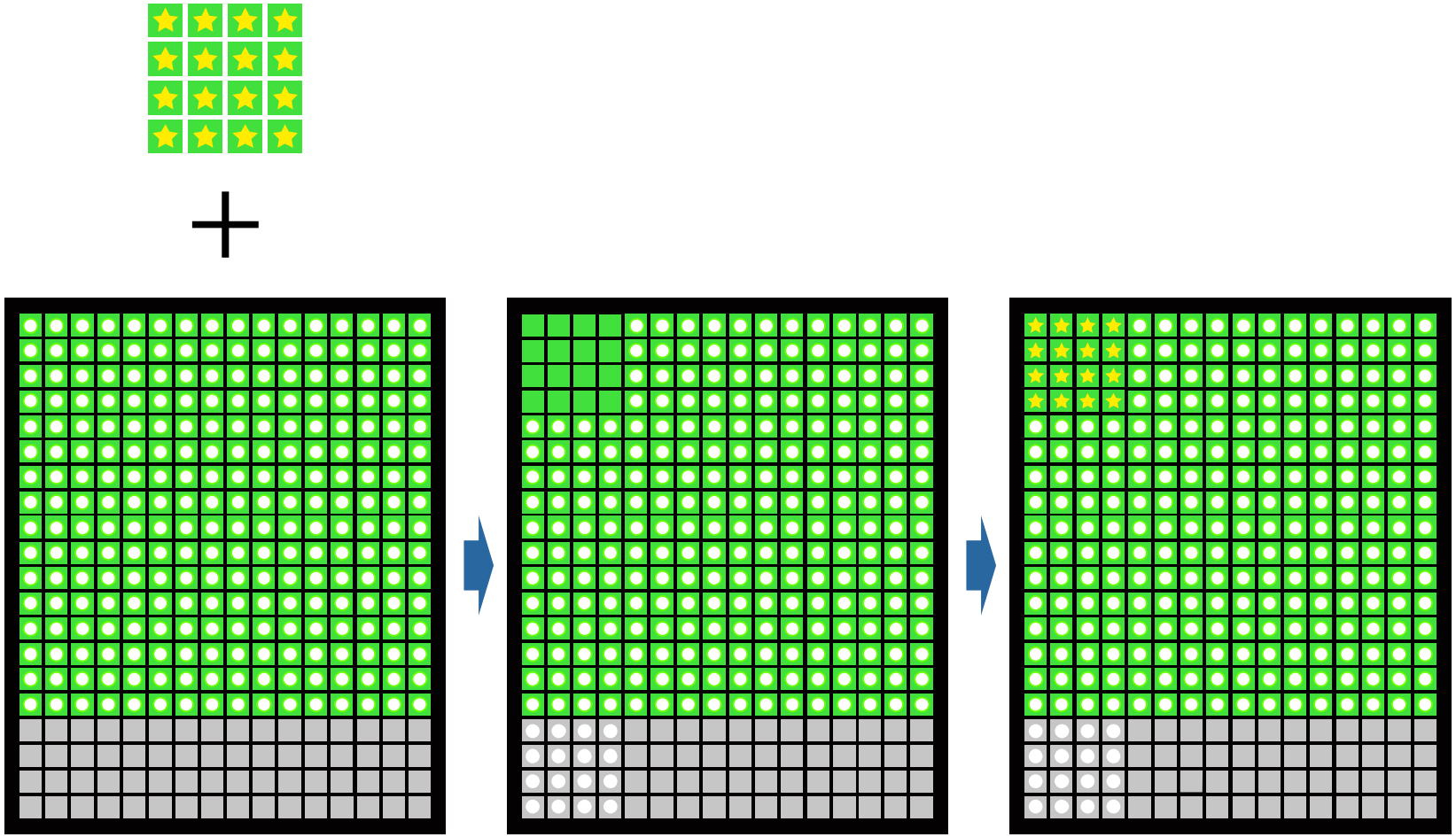




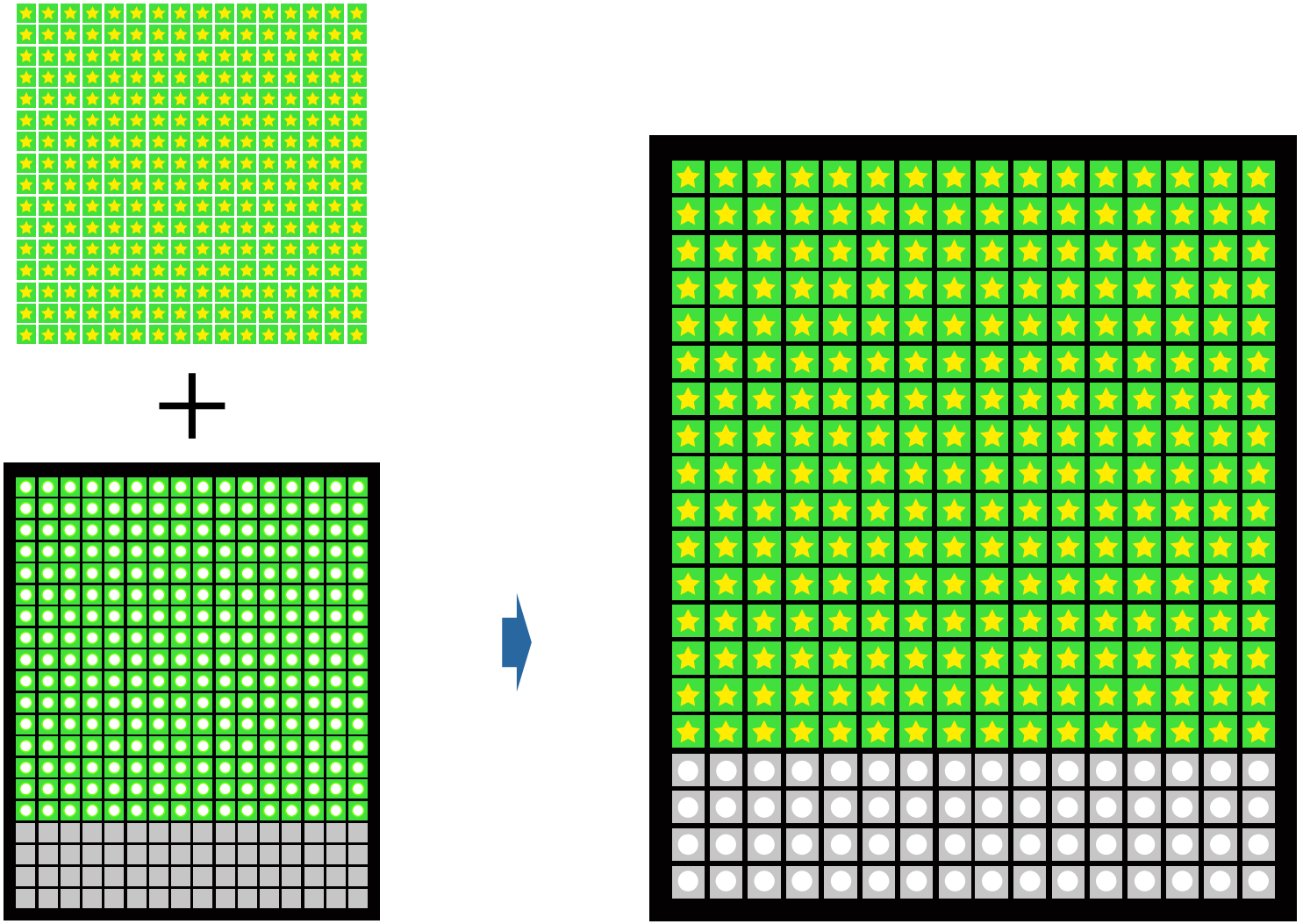
SSDのデータ書き込み



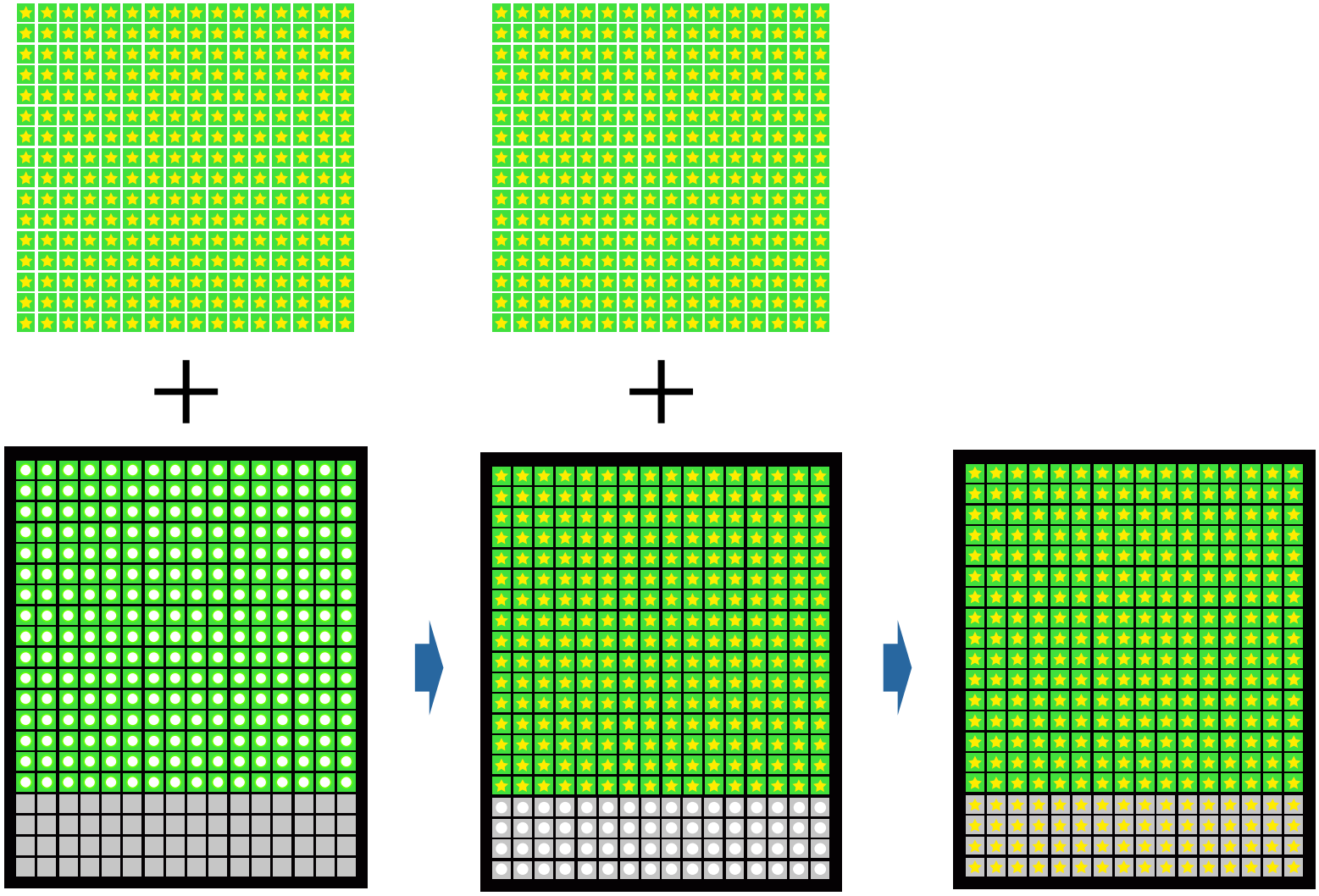
SSDのデータ書き込み



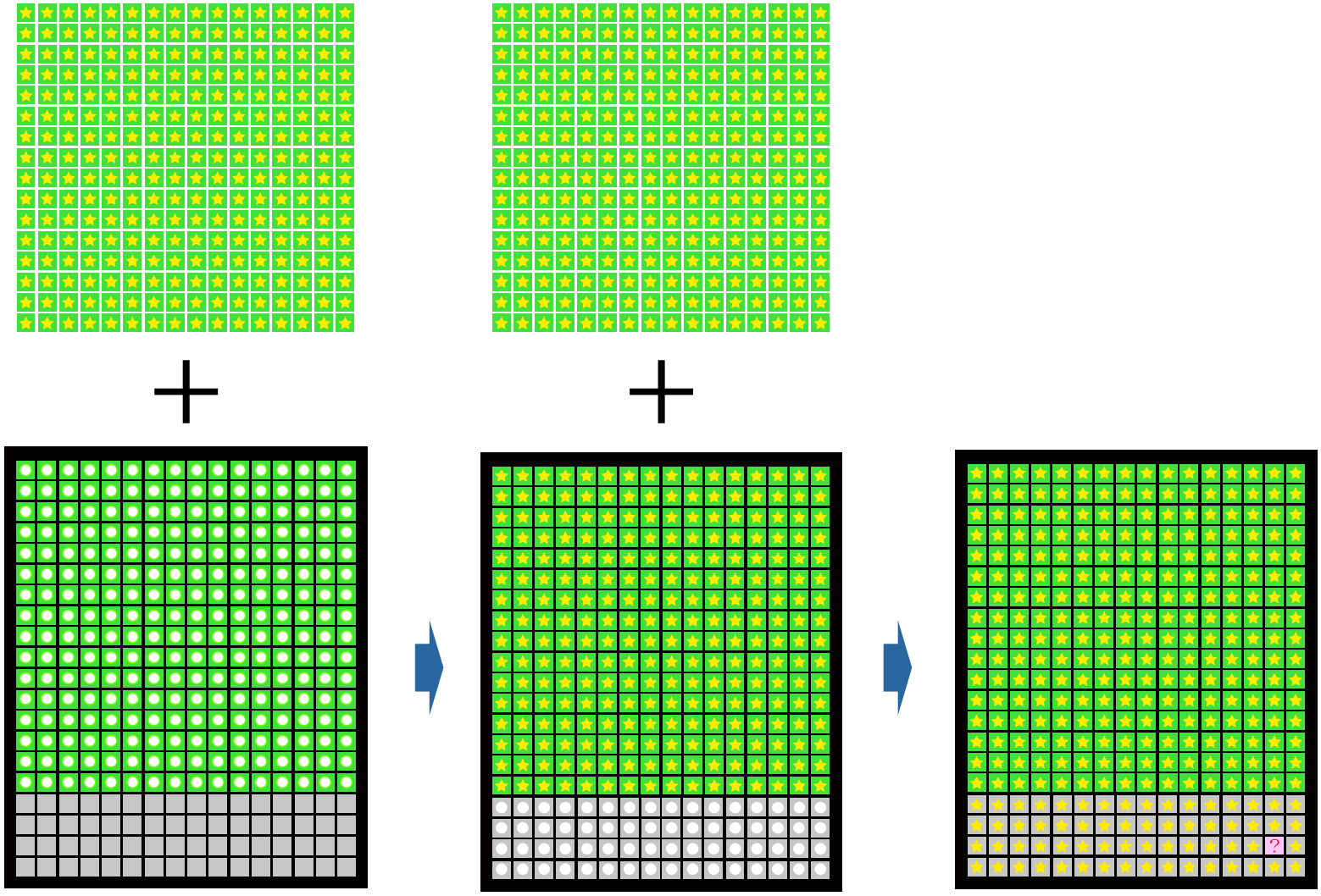
SSDのデータ書き込み



SSDのデータ書き込み



SSDのデータ書き込み



データ消去は、決して簡単な話ではない

○ADEC消去技術認証　認証数／検証数について

◆総認証数推移

	2018	2019	2020	総数
認証数	4	6	4	14

※2019年はOEM1件含

◆内訳（予備判定／実検証対応）

■2018年（メディア別検証結果）
申請1：HDD〇　SSD〇
申請2：HDD〇　SSD〇
申請3：HDD〇　SSD〇
申請4：HDD〇　SSD×→再検証時〇

■2020年（メディア別検証結果）
申請1：HDD〇　SSD〇
申請2：HDD〇　SSD〇
申請3：HDD〇　SSD〇
申請4：HDD〇　SSD×→再検証予定

■2019年（メディア別検証結果）
申請1：HDD〇　SSD×→再検証時〇
申請2：HDD〇
申請3：HDD×→再検証無し（2台検証）
申請4：HDD〇　SSD〇
申請5：HDD〇　SSD〇
申請6：HDD〇　SSD〇

2020年11月　ADEC（データ適正消去実行証明協議会）提供資料

データ消去は、決して簡単な話ではない

○ADEC消去技術認証　認証数／検証数について

◆内訳（予備判定／実検証対応：メディア、消去ランク別）

		2018		2019		2020		計
総検証数		8		9		8		25
		適合	不適合	適合	不適合	適合	不適合	
HDD	Clear	4	0	3	3	1	0	11
	Purge	0	0	0	0	3	0	3
SSD	Clear	3	1	2	1	1	0	9
	Purge	0	0	0	0	2	1	3
計		7	1	5	4	7	1	

■総検証数25件
○メディア／消去ランク別総数
・HDD：14件
＜適合：11件（Clear9／Purge2）、不適合：3件（Clear2／Purge1）＞
・SSD：11件
＜適合：9件（Clear6／Purge3）、不適合：2件（Clear2／Purge0）＞

※2020年よりPurge検証を採用

2020年11月　ADEC（データ適正消去実行証明協議会）提供資料

参考資料（日本語）

「証拠保全先媒体のデータ抹消に関する報告書」 デジタル・フォレンジック研究会
https://digitalforensic.jp/home/act/products/data_report/

「日本のPCリユースにおけるデータ消去について」 伊藤 修司
https://digitalforensic.jp/wp-content/uploads/2016/02/pc-reuse.pdf

「データ抹消に関する米国文書（規格）及び HDD、SSD の技術解説」 沼田 理
https://digitalforensic.jp/wp-content/uploads/2016/02/technical-aspect.pdf

「消去アクセス難易度別にみるHDDのデータ領域3分類」下垣内 太
https://digitalforensic.jp/wp-content/uploads/2016/04/3-categories-of-hdd-data.pdf

「データ消去に関する海外規格の動向」 瀧澤 和子
https://digitalforensic.jp/wp-content/uploads/2016/02/standards.pdf

「データ抹消に関する実態調査」 宇野 幸治
https://digitalforensic.jp/wp-content/uploads/2016/02/survey.pdf

「データ抹消に関する性能評価報告」 宇野 幸治、土井 洋
https://digitalforensic.jp/wp-content/uploads/2016/02/performance-evaluation-.pdf

「デジタル・フォレンジックの有効性
—セキュリティマネジメントからみたPCデータ抹消について—」 山口 大輔
https://digitalforensic.jp/wp-content/uploads/2016/04/management.pdf

---

「データ消去技術 ガイドブック」 データ適正消去実行証明協議会（ADEC）
https://adec-cert.jp/guidebook/index.html

「CODE BLUE 2016, EXOTIC DATA RECOVERY & PARADAIS」 Dai Shimogaito
https://www.youtube.com/watch?v=lIJkFJTFKLY

講演内容に関するご質問など

ご聴講ありがとうございました。

本日の解説内容について、ご質問などございます場合には、
下記までお問い合わせください。

また、不足している点や、誤っている情報などございましたら、
ぜひ教えていただけるとうれしいです。

下垣内 太 （ しもがいと だい ）
dai.shimogaito@gmail.com
https://www.facebook.com/dai.shimogaito