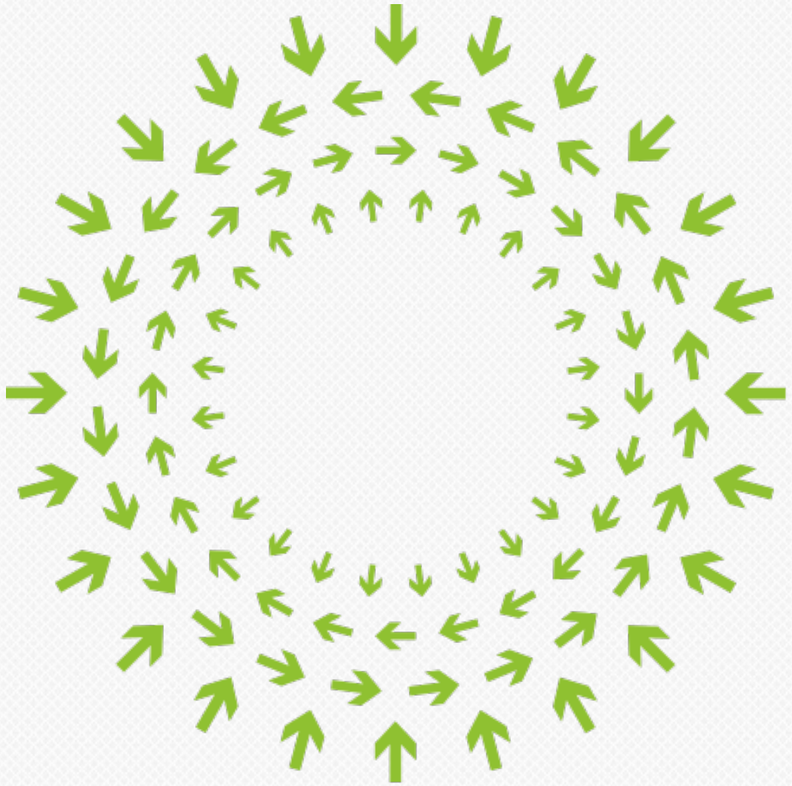


OPEN

Compute Engineering Workshop

March 9, 2015

San Jose

OPEN
Compute Engineering Workshop

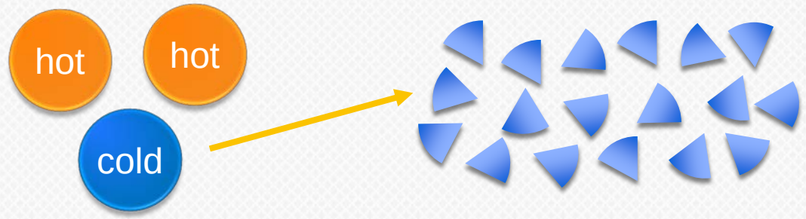
“Igloo” Cold Storage Concept

Keeping your cool data safe and available for a long time to come

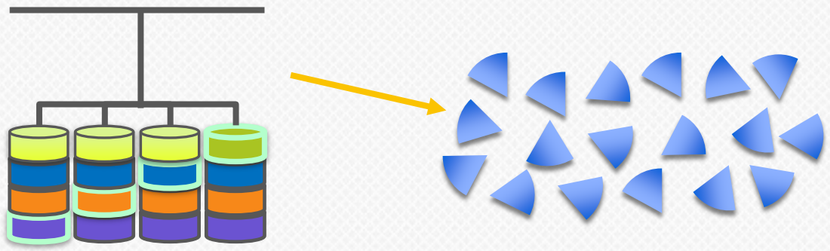
Rob Ryan
WD Labs™
Scientist

Cold Storage Premise

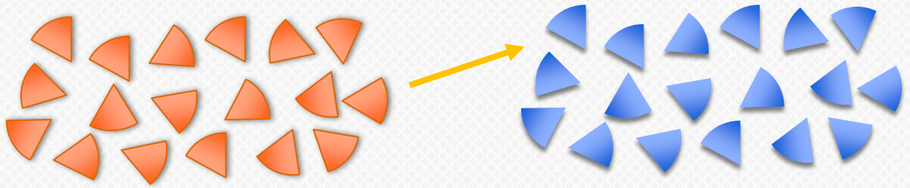
Replication use case could move the third replica to cold



Move Traditional Raid data to an erasure coded Archive as data cools



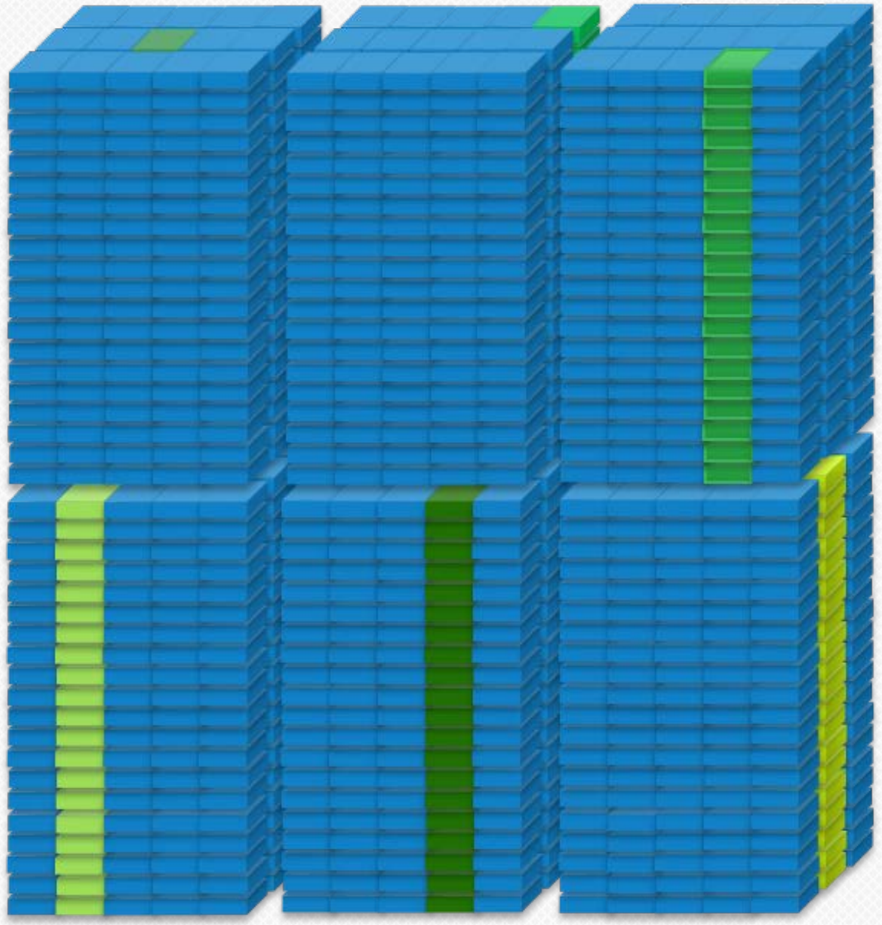
Move LRC code data to Archive



Deep Cold Storage

Power Zones can save massive amounts of power

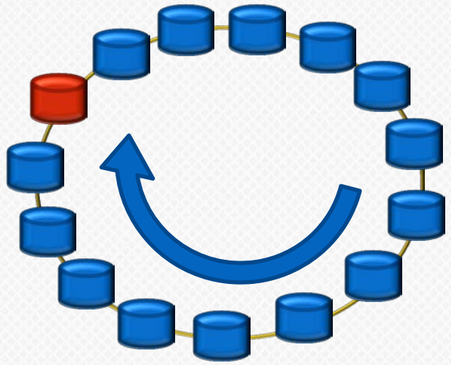
“Write instantly” available utilizing any active power zone



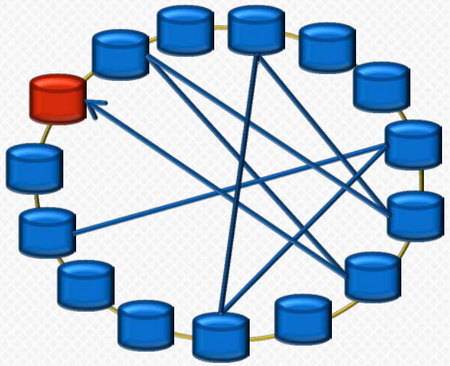
“Read seldom” allows recovery in minutes

Read Options

Scheduled “power zone” or “on demand” queue options



Predictable
10 min stops
75 minute average latency

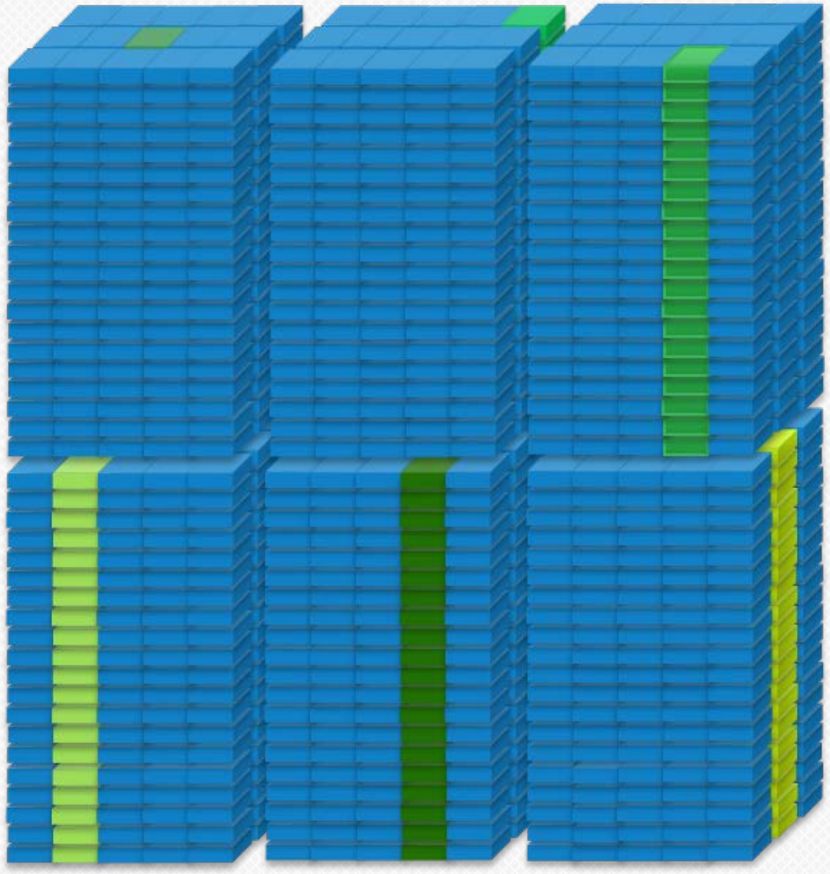


Travelling Salesman
Stops queued by “values”
Few minute or more latency
depending on load and algorithm

Conceptual Example

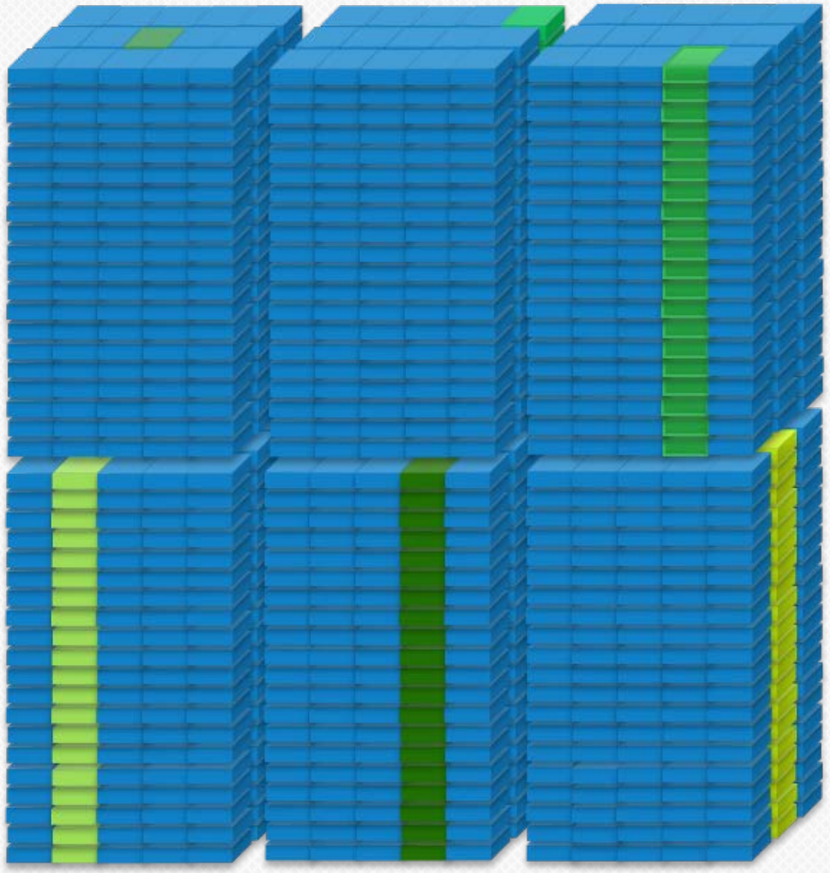
Triple Rack

1800 Drives
10.8PB
120 Drives Active
1200W



- 6 Power Groups
- Power Group contains 20 active trays
- Each Tray in a Group contains 15 power zones
- Group switch time may be staggered for 100% write availability
- 20 active drives in each group allow erasure coding to be spread with some over provisioning in case of failure
- Actual algorithm would scatter these neat groups to many random racks

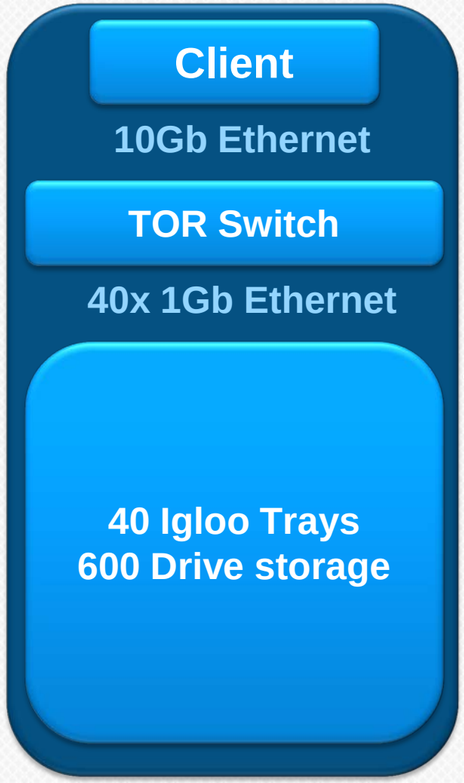
Network Attached Deep Cold Storage Tray



Traditional
Cold Storage

5360W or 0.62W/TB
@480 6TB/Drive

Ultra low power
Cold Storage



1160W or 0.13W/TB
@480 6TB/Drive

Triple Rack
120x 1GB TOR
120 trays
120 compute nodes
Network Attached
10.8PB
1200W
+200W switch

0.13 W/TB

Network Attached Deep Cold Storage Tray

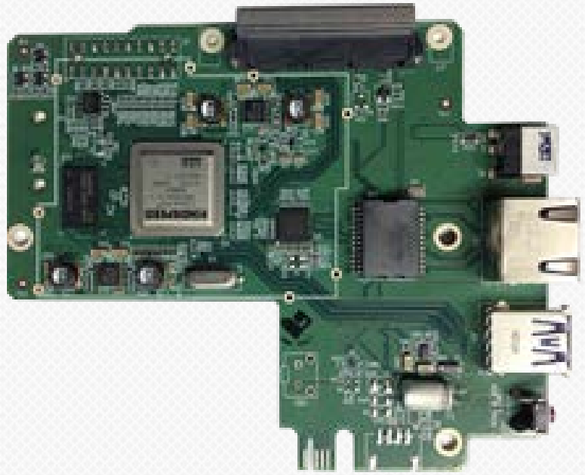
HDA

x15

HDD PCBA

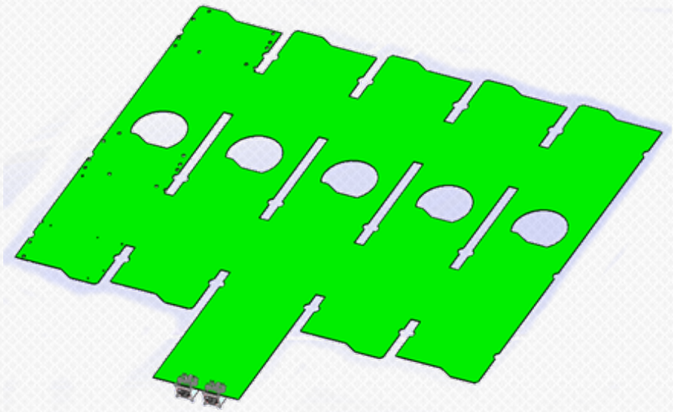
x15

WD NAS HW

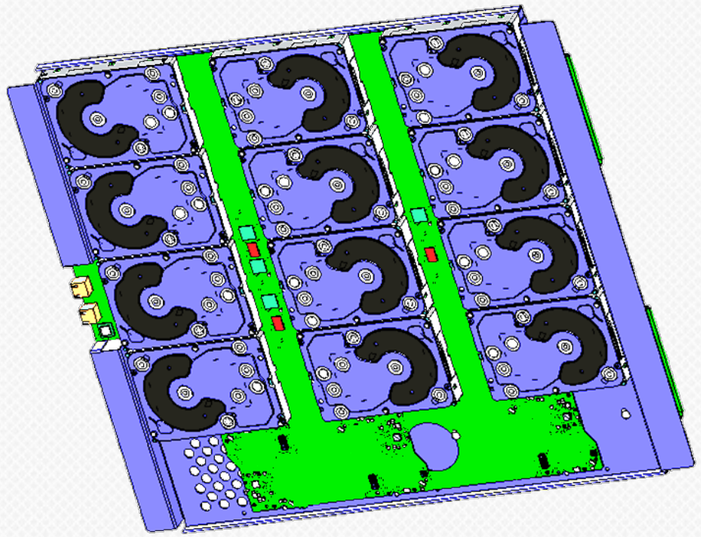


Network Attached Deep Cold Storage Tray

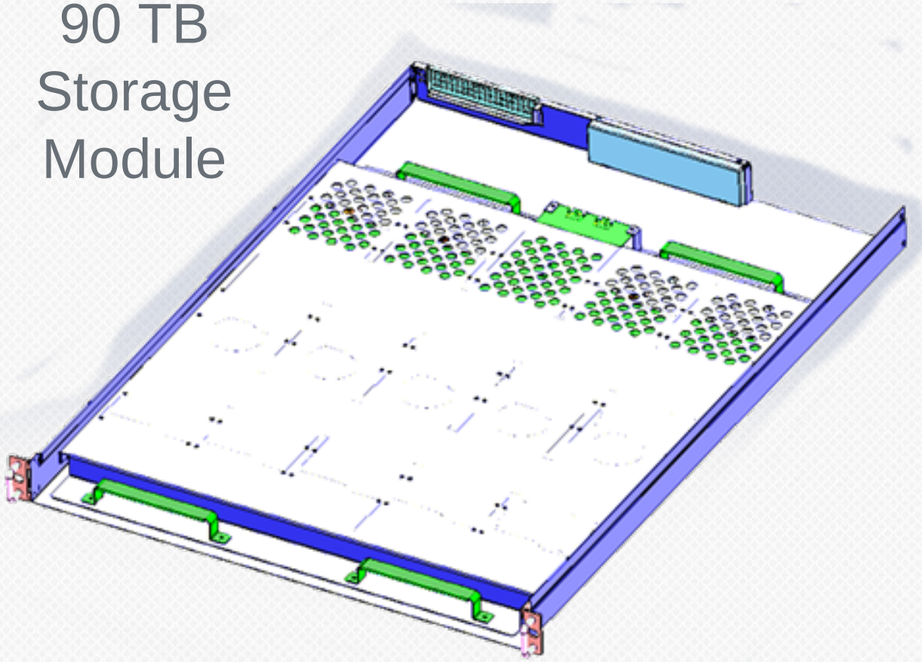
Single PCBA



Bottom Mount the HDAs



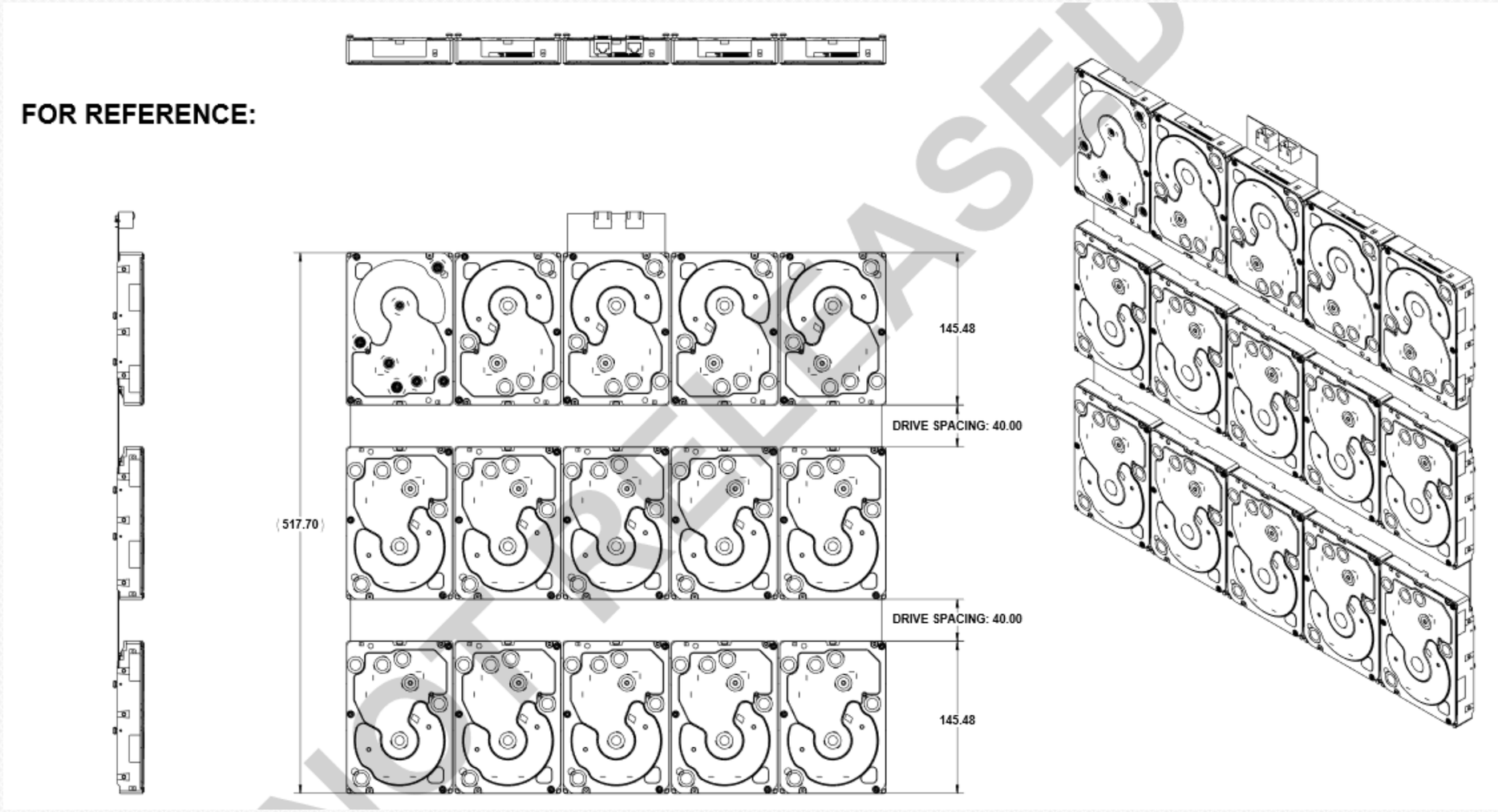
90 TB Storage Module



10W power consumption for the whole Tray

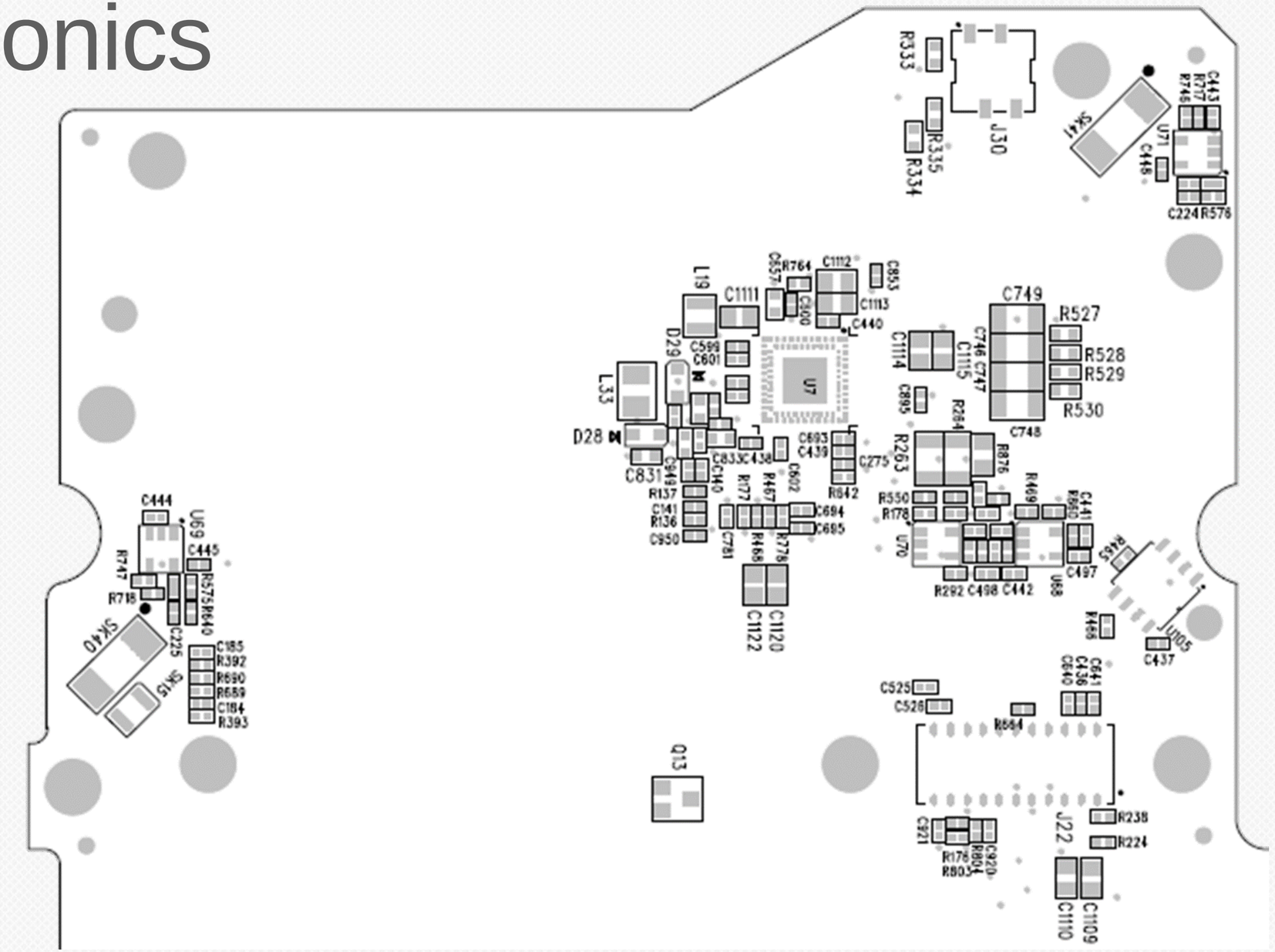
Cold Storage platform provisioned for single drive function per group

PCBA Mech Assembly



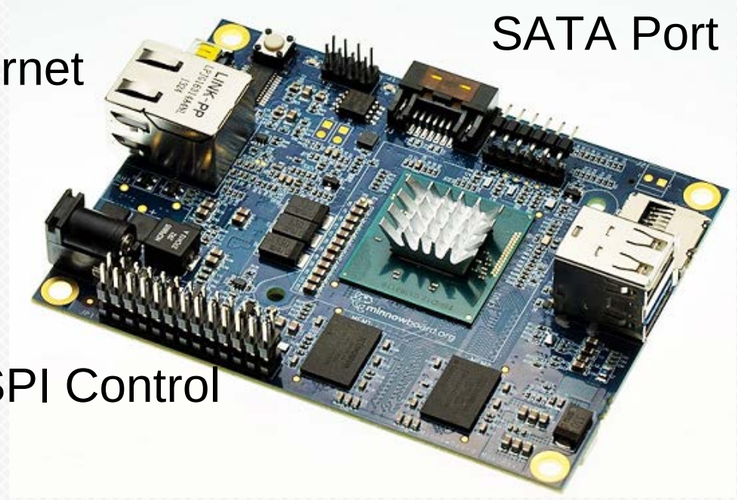
HDA electronics

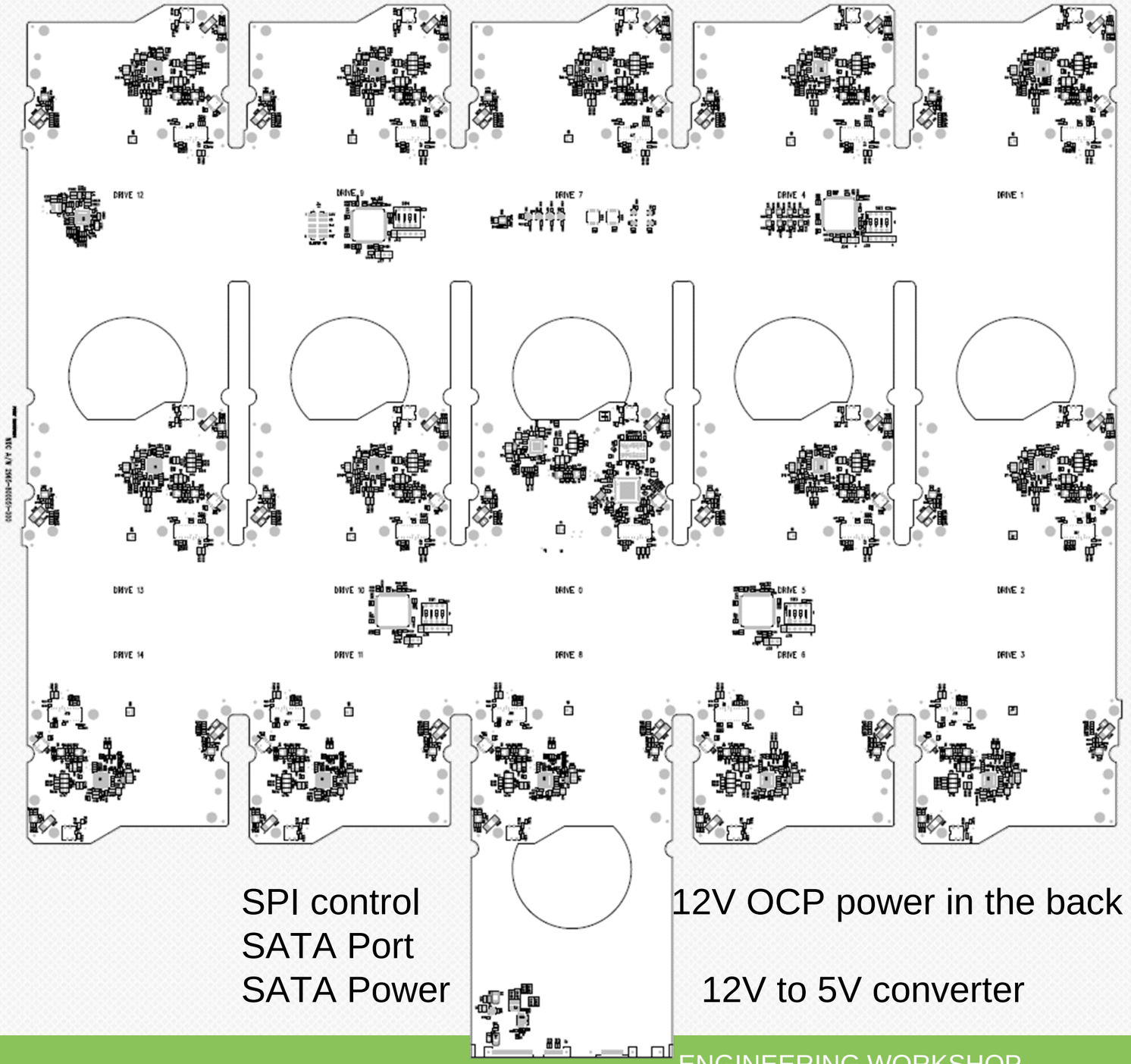
- Motor Control
- Sensors
- Connectors



Total System

- Single SoC
- Muxes
- HDA support
- Computing Node





Thermal

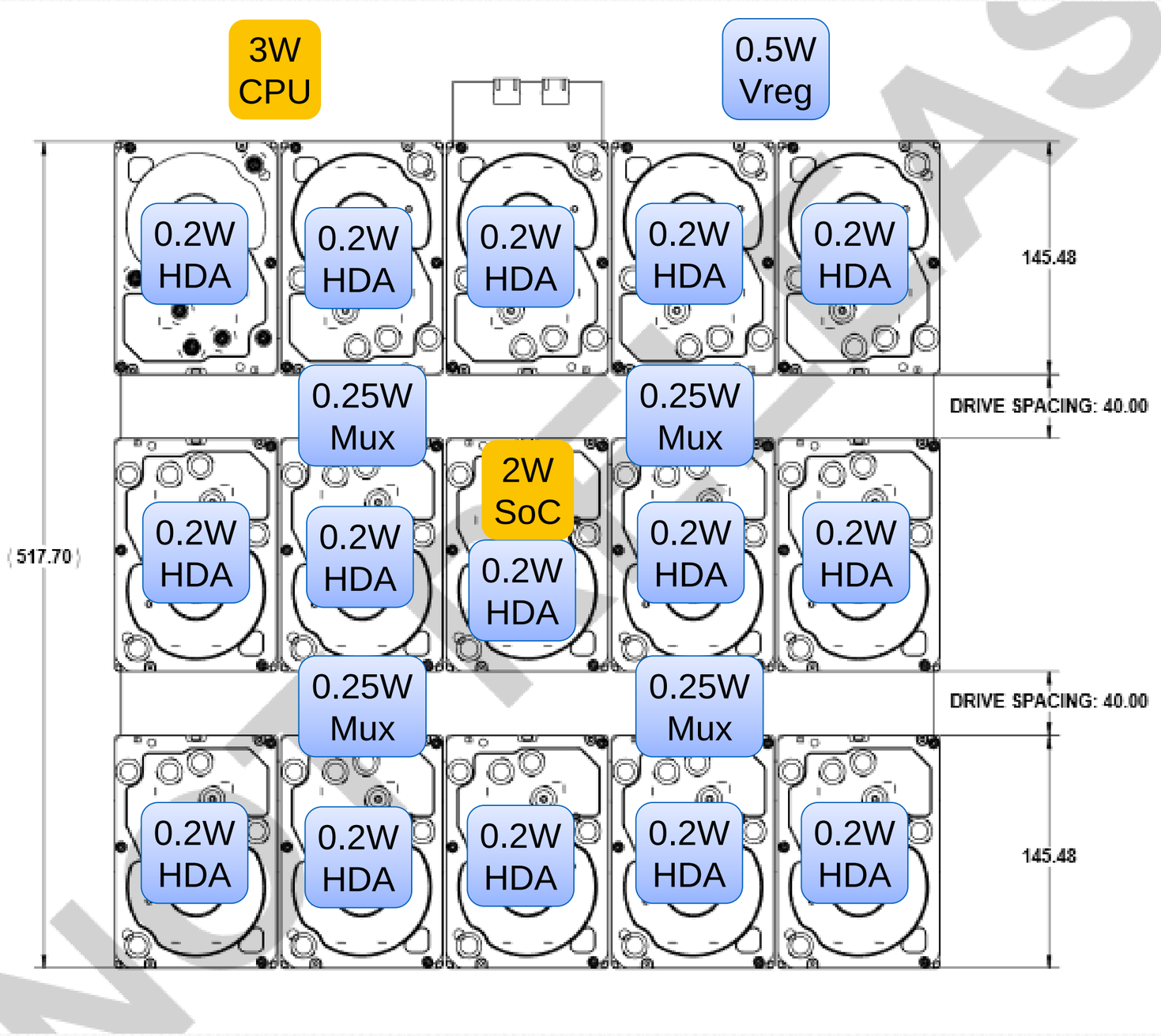
- Thermal time constant of a 3.5” HDD is about an hour.
- If each HDA is only enabled 7% of the time, and the “On time” is 1/6 of the time constant, the HDD will not get close to maximum operating temperature.
- It is conceivable that a 1U Pizza Box dissipating 10W would not need a fan.
  - Napkin calculation shows about 10C rise for 10W heater with front and back of rack unit exposed to air

Thermal Map

- 3W CPU
- 2W SoC
- 0.5W 12V to 5V reg efficiency
- 15x 0.2W=3W HDAs
- 4x 0.25W=1W Mux power

- 9.5W total

- HDA is 3W for 10 min and then 0
- Spin up is 20W for 5 seconds



Q&A

- How compelling is the Deep Cold use case?
- Is the “WD Igloo” ease of use appealing?

“Igloo” Deep Cold Storage Concept

- In development at WD Labs™

Thank you

Backup Power

6.1 Rack Power Budgets

OCP Cold Storage spec v0.7

Estimated power consumption for a Cold Storage rack is:

- Storage unit (Open Vault with only 2 HDD spinning): 70W
- Compute node: 300W
- Network switch: 200W
- Power budget without network switch: 70 x 16 + 300 x 2 = 1,720W
- Power budget with network switch: 70 x 16 + 300 x 2 + 200 = 1,920W
- Power budget for every three racks with one network switch: 1,720 x 2 + 1,920 = 5,360W

- Igloo with 1 HDD spinning 10W vs 35W
- Compute node 0 vs 300W
- Network Switch 200W
- Power budget without network switch10x32 = 320W
- Power budget with network switch 10x32+200=520W
- Power budget for every three racks with one network switch 320x2+520=1160W