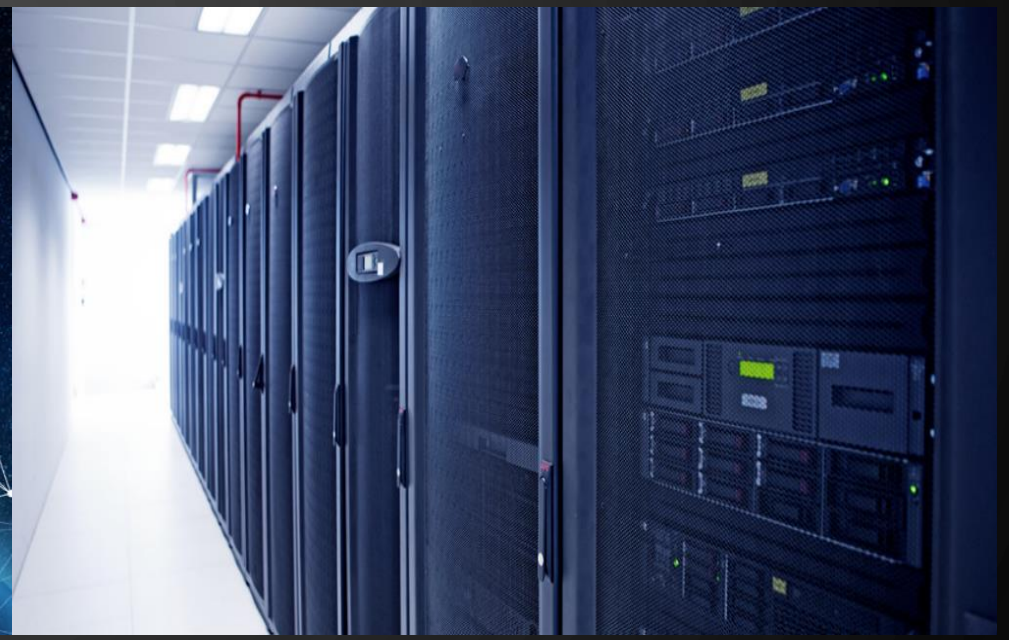
PRESERVATION OR DELETION: ARCHIVING AND ACCESSING THE DATAVERSE

- **In 2030…**
  - **The hot (nanoseconds-to-milliseconds access) layers will comprise 8% of enterprise data…**
  - **The warm (milliseconds-to-seconds access) layers will comprise 17% of enterprise data…**
  - **The cool (minutes-to-24 hours access) layers will comprise 14% of enterprise data…**
  - **The cold (days-to-1 week access) layers will comprise 26% of enterprise data…**
  - **The frozen (weeks-to-years access) layers will comprise 35% of enterprise data…**
- **In 2030, the effective "active archive" will comprise 60% to 75% of enterprise data…**

John Monroe
torstones@gmail.com
Mobile Office: +1 (408) 828-3967

Furthur Market Research

THE EXPANDING DATAVERSE



- The enterprise “data pools” of the early 2000s became “data lakes” by 2010 and grew in recent years to become “data oceans” which have already begun to morph into a vast multiform “dataverse.” And because “data is the new oil,” we are loathe to delete any data.
- Increased storage at any point in the World Wide Web—bear in mind that a mobile phone is a point in the Web—increases the possibilities for storage in every part of the Web.
- We are only beginning to see the enormous implications of that simple fact.

AGENDA

- New Forecasts
- New Delineations of Enterprise Data
- Recent Surveys
- Inconclusive Conclusions

Enduring Question: *Will the Past be Prologue, or Will History Be Bunk?*

Note: My forecasts are always devised with these precautionary adages in mind:

- *The only thing we know with certainty about any forecast is that it will be wrong. — Anonymous*
- *He who foretells the future lies, even if he tells the truth. —Arab Proverb*

FORECAST OF ENTERPRISE PETABYTES DELIVERED 2023-2030 AND ACTIVE INSTALLED BASE ESTIMATES

	2010	2015	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	"CAGR 2023-2030"
Enterprise SSD PB	187	26,154	130,766	178,972	186,588	230,664	334,195	498,792	690,849	925,047	1,098,030	1,604,222	2,050,196	36.6
Annual Growth %	–	223.9	64.1	36.9	4.3	23.6	44.9	49.3	38.5	33.9	18.7	46.1	27.8	
Enterprise HDD PB	45,216	157,093	679,887	959,011	941,749	1,251,584	1,610,789	2,177,786	2,735,300	3,605,125	4,589,324	6,021,194	8,062,378	30.5
Annual Growth %	–	34.7	39.9	41.1	–1.8	32.9	28.7	35.2	25.6	31.8	27.3	31.2	33.9	
Enterprise Tape PB	30,208	98,432	136,119	189,938	206,842	259,794	303,440	366,252	464,773	574,460	763,457	966,537	1,258,431	25.3
Annual Growth %	–	15.0	1.4	39.5	8.9	25.6	16.8	20.7	26.9	23.6	32.9	26.6	30.2	
Total Compressed Shipments PB	75,611	281,679	946,772	1,327,921	1,335,179	1,742,043	2,248,424	3,042,830	3,890,922	5,104,631	6,450,812	8,591,952	11,371,005	30.7
Annual Growth %	–	33.9	33.6	40.3	0.5	30.5	29.1	35.3	27.9	31.2	26.4	33.2	32.3	
Active Installed Base PB	**91,000**	**819,949**	**2,923,201**	**3,950,945**	**5,232,405**	**6,447,587**	**7,985,007**	**10,081,065**	**12,644,067**	**16,413,518**	**21,122,288**	**27,465,816**	**35,793,990**	

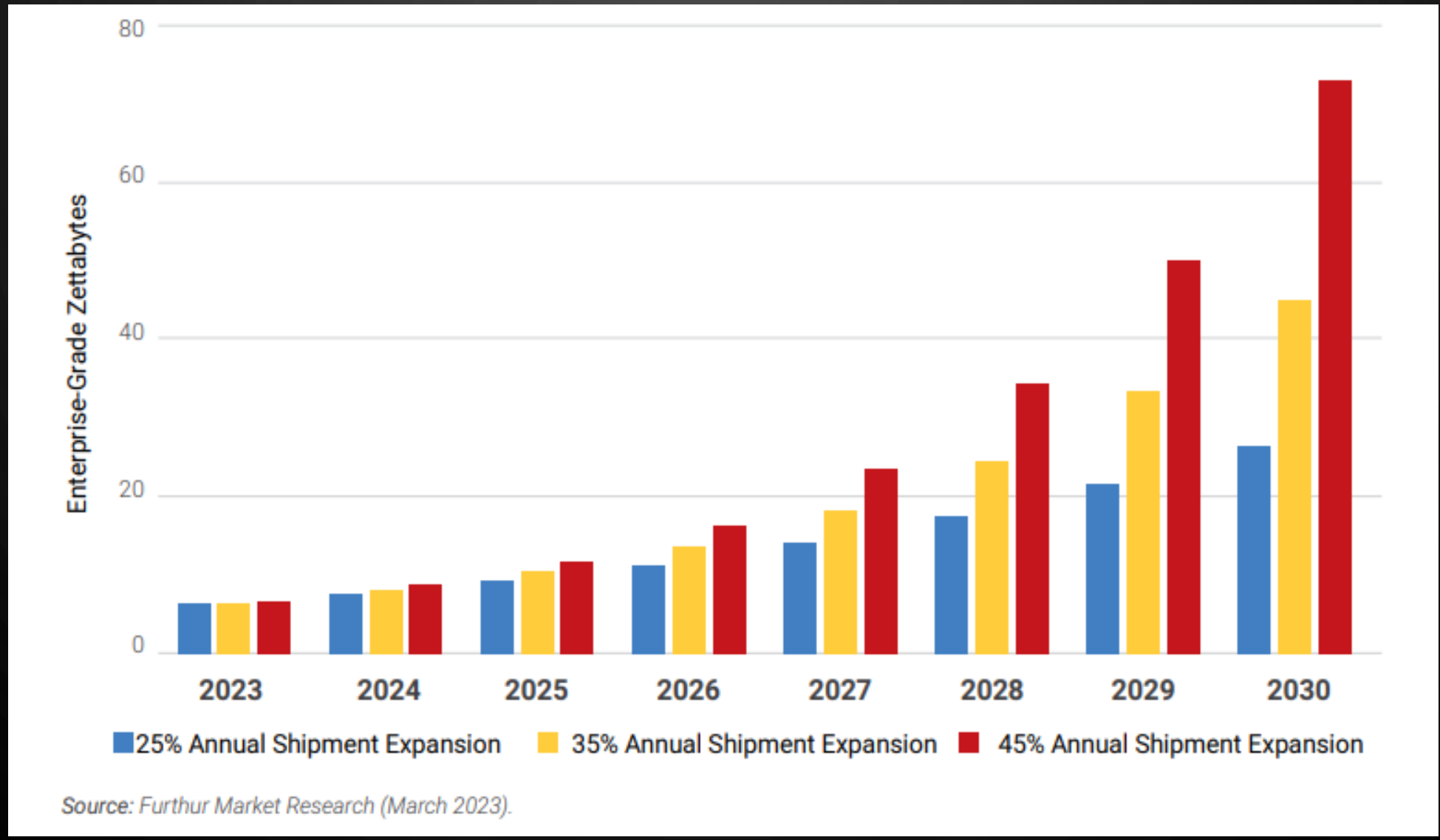
Source: Furthur Market Research (March 2023)

ALTERNATE 2023-2030 GROWTH SCENARIOS

	2023	2024	2025	2026	2027	2028	2029	2030
Alternate 2023-2030 Shipment Scenarios								
Total Shipped Enterprise PB Expanding at 25%/Year 2023-2030	1,668,974	2,086,218	2,607,772	3,259,716	4,074,644	5,093,306	6,366,632	7,958,290
Total Shipped Enterprise PB Expanding at 35%/Year 2023-2030	1,802,492	2,433,365	3,285,042	4,434,807	5,986,990	8,082,436	10,911,288	14,730,239
Total Shipped Enterprise PB Expanding at 45%/Year 2023-2030	1,936,010	2,807,215	4,070,462	5,902,169	8,558,145	12,409,311	17,993,501	26,090,576
Alternate 2023-2030 Active Installed Base Scenarios								
Active Installed Base PB at 25% Annual Shipment Expansion	6,374,518	7,749,733	9,410,733	11,342,528	14,081,993	17,433,256	21,551,464	26,466,924
Active Installed Base PB at 35% Annual Shipment Expansion	6,508,036	8,230,398	10,568,668	13,675,554	18,327,364	24,667,757	33,330,622	45,018,031
Active Installed Base PB at 45% Annual Shipment Expansion	6,641,554	8,737,766	11,861,455	16,435,704	23,658,670	34,325,938	50,071,015	73,118,761

Source: Furthur Market Research (March 2023).

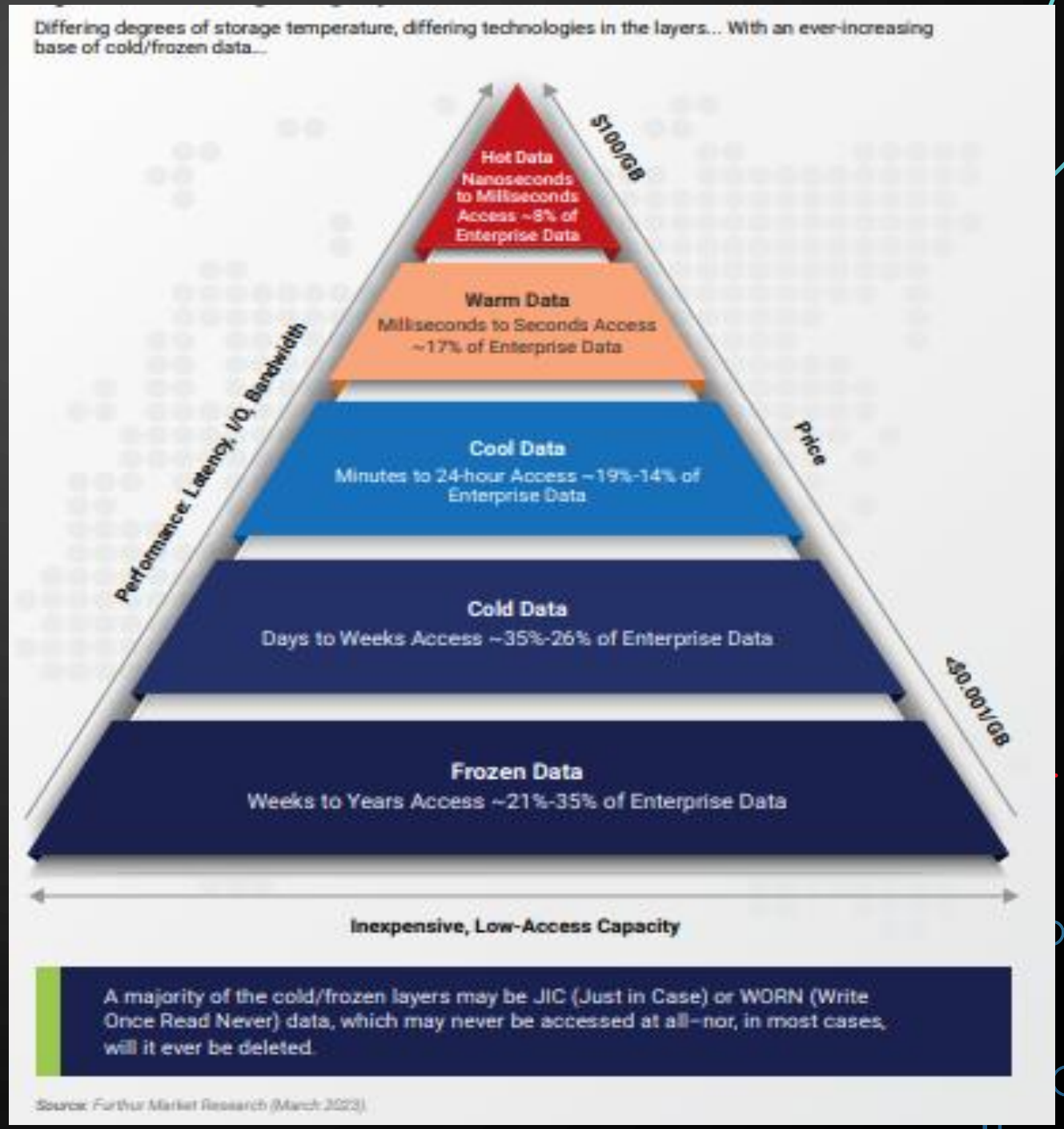
ALTERNATE 2023-2030 GROWTH SCENARIOS



Source: Furthur Market Research (March 2023).

NEW DELINEATIONS OF ENTERPRISE DATA BASED ON ESTIMATED ACCESS FREQUENCY

- The hot and warm data layers will remain fairly constant at ~25% of the total (8% hot, 17% warm).
- The cool data layer shrinks somewhat from ~19% in 2020 to ~14% in 2030 and the cold data layer shrinks as well, from 35% in 2020 to 26% in 2030, while the frozen data layer grows from ~21% in 2020 to ~35% in 2030.
- With the advent of more strictly enforced AI/ML corporate archive and access rules, it is also possible that, in many data centers, 25% of the data will grow to be hot in 2030, while the warm and cool data layers decline to insignificance, and the cold/frozen data layers will grow to 75% of the total.
- **In 2030, the cool/cold layers will comprise 40% of active installed based of enterprise data (14ZB-17ZB) and the cold/frozen layers will comprise 61% (~22ZB-26ZB) of the active installed base of enterprise data.**



EVOLVING ACTIVE ARCHIVE REVENUE OPPORTUNITIES

	2023	2024	2025	2026	2027	2028	2029	2030
Evolving Opportunities to Manage Cold/"Frozen" Data, Alternate Scenarios								
57.4%-61.2% ofTotal Shipped PB Expanding at 25%/Year 2023-2030	957,991	1,212,093	1,530,762	1,926,492	2,424,413	3,050,890	3,826,346	4,790,891
Vendor Revenue Opportunity In Billions of USD	$4.9	$4.7	$5.1	$5.4	$5.6	$5.8	$6.0	$6.2
57.4%-61.2% ofTotal Shipped PB Expanding at 35%/Year 2023-2030	1,034,631	1,413,785	1,928,320	2,620,971	3,562,259	4,841,379	6,557,684	8,867,604
Vendor Revenue Opportunity In Billions of USD	$5.3	$5.5	$6.4	$7.3	$8.2	$9.2	$10.3	$11.5
57.4%-61.2% ofTotal Shipped PB Expanding at 45%/Year 2023-2030	1,111,270	1,630,992	2,389,361	3,488,182	5,092,097	7,433,177	10,814,094	15,706,527
Vendor Revenue Opportunity In Billions of USD	$5.7	$6.4	$8.0	$9.7	$11.7	$14.0	$17.0	$20.4
Aggressive Best Ex-Factory ASPs/GB for Tape or Other Technology	**$0.0051**	**$0.0039**	**$0.0033**	**$0.0028**	**$0.0023**	**$0.0019**	**$0.0016**	**$0.0013**

Source: Furthur Market Research (March 2023)

RECENT SURVEY RESULTS

Interviews with IT managers of 50PB-500PB databases revealed the following:

- Almost all of these data center managers specified “indefinite” retention periods for the vast majority of their data, but they fear that the rising sustainability costs of preserving their data for many years or “indefinitely” will become prohibitive.
- For all these managers, data immutability was a crucial issue—all aspects of the original data absolutely must remain unchanged.
- Many of these IT managers classified their data as 100% “cold,” but it could become “hot” at any time depending on data access requests —in other words, 100% of their data is an “active archive.”
- Several enterprise IT managers with whom we spoke stated that an exacerbated problem with data deletion is establishing generally agreed-upon ground rules. When they asked for buy-in from their internal clients, they could not obtain any solid commitment for, say, 5-year, 7-year, or 10-year deletion objectives for aging data. There was always the lingering fear that after 5 years or 7 years or 10 years and 1 day, they would absolutely need that old data for some unspecified, but critical, future purpose.

INCONCLUSIVE CONCLUSIONS

- The data centers of the future will need everything the SSD, HDD, and tape industries can manufacture and deliver, as well as requiring new DNA and optical and perhaps other enterprise storage technologies, to cost-effectively and reliably preserve the priceless artifacts of our personal, corporate, and cultural history.
- Availability and sustainability challenges will create a global need for "autonomic" data systems that can provide intelligent "active archive" management and seamless migrations of hot-to-warm-to-cool-to-cold-to-frozen data and back again, from core to edge to cloud.

INCONCLUSIVE CONCLUSIONS

- The costs of managing our multi-millionfold-petabyte dataverse over increasingly lengthy time periods will create new use cases for old storage technologies and demand the creation of new, more cost-effective, and power-efficient storage technologies.
- *Inevitably and inescapably, richly varied computing technologies will come and go, but the DATA we create will remain, and will grow to unimaginable immensity.*

The Dataverse…

An enlargement of the library of forms in which DATA, unleashed in fresh dimensions, can come to profitable life…