

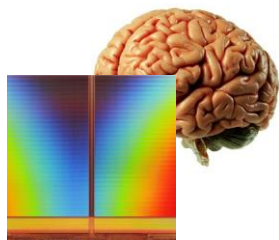


Flash Memory Summit
2018

Tachyum™ Prodigy



Dr. Radoslav Danilak



Legal Disclaimers

NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. TACHYUM ASSUMES NO LIABILITY WHATSOEVER AND DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF TACHYUM PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, OR MERCHANTABILITY.

All information provided here is subject to change without notice. Nothing in these materials is an offer to sell any of the components or devices referenced herein.

Tachyum is a trademark of Tachyum Ltd., registered in the United States and other countries, Tachyum Prodigy is a trademarks of Tachyum Ltd. Other products and brand names may be trademarks or registered trademarks of their respective owners.

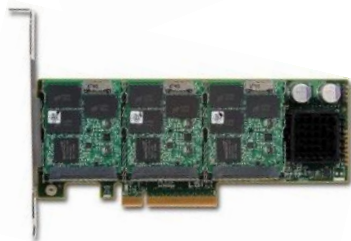
©2018 Tachyum Ltd. All Rights Reserved.



Flash Memory Summit
2018

10 Years of Leading World-Class Innovation

SandForce™



10x Flash Life
\$20 → \$3 / GB
SLC → MLC

skyera



100x Flash Life
\$20 → \$3 → \$1 / GB
eMLC → MLC → TLC
Compression + Dedup.

Tachyum™



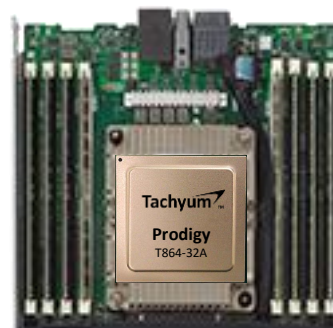
300x Flash Life
25¢ → 9¢ → **1¢ / GB**
TLC → QLC
Compression + Dedup.
Hyperscale-Out

Flash-Only Datacenter for Lower Cost & Power

- Flash is already cheaper than 10TB disk drive in hyperscale/Hadoop system
 - Disk 11¢/GB: 3 copies 10TB 3.5" \$320 HDD = 9.6¢/GB + 1.4¢/GB system
 - Flash 9¢/GB: DRAMeXchange 32GB USB \$2.5 mCOB = 7.8¢/GB + 1.2¢/GB system
- 1¢/GB effective achievable for flash
 - 5:1 compression + deduplication, 2:1 thin provisioning, zero overhead snapshots + clones
- 3 copies vs. RAID6 used to avoid 4x slowdown of slow HDD & high CPU cost
 - RAID6 for write requires 3 reads and 3 writes reduces 2x performance at 4:1 read/write ratio
 - If drive is failed then for 1/n drives (n-1) drives needs to be read leading to 2x slow down



8 x 64-256GB RDIMM



2 x 400G Ethernet

4 DDR5 200GB/s
500GB/s HBM



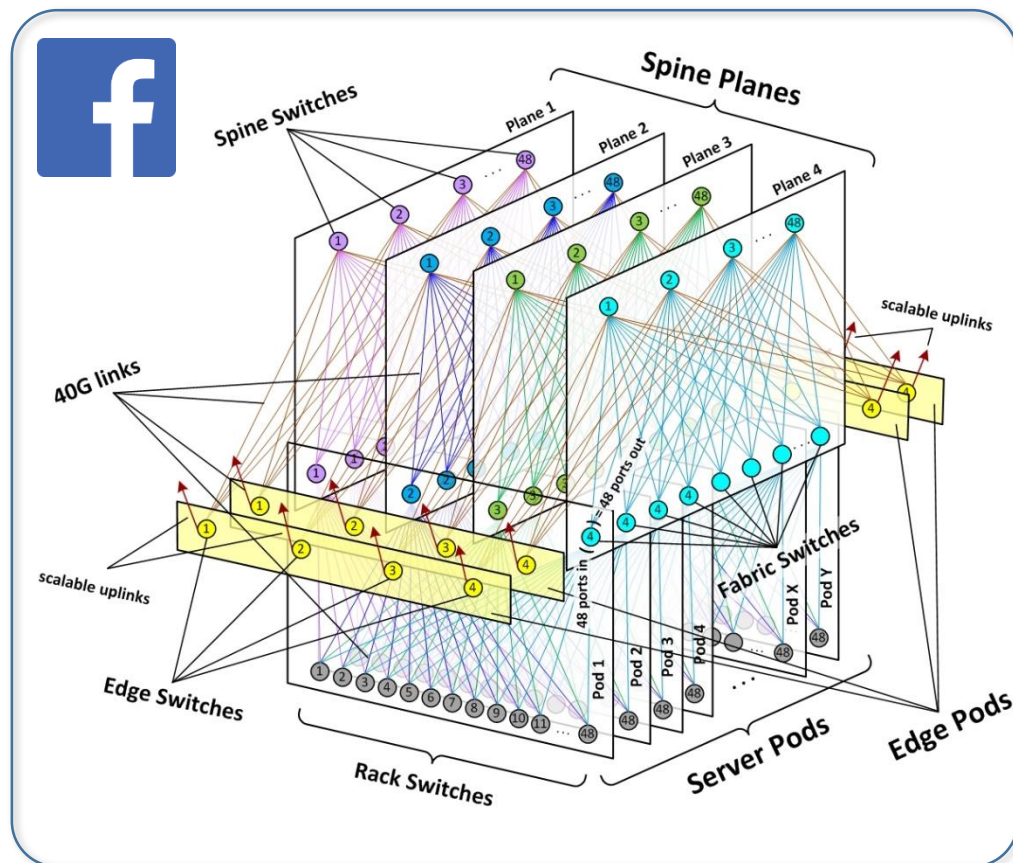
4 DDR5 200GB/s
500GB/s HBM

32 x PCIE 5.0 x2
500GB/s SSDs



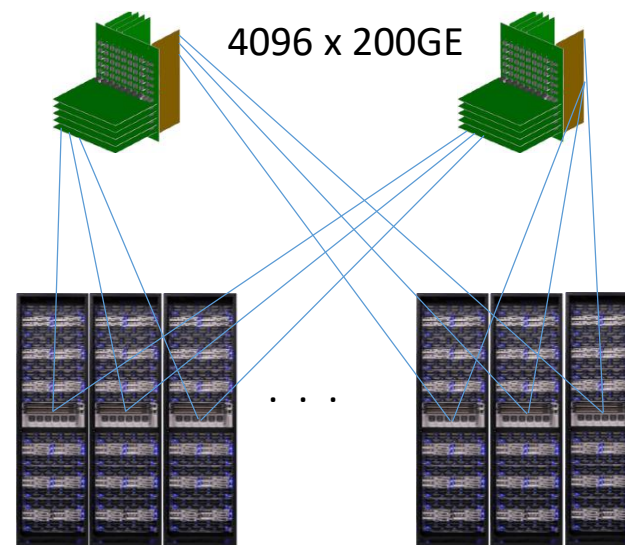
Networking: 10x Bandwidth at Same Cost

Copper Rack → Edge → Fabric → Spine

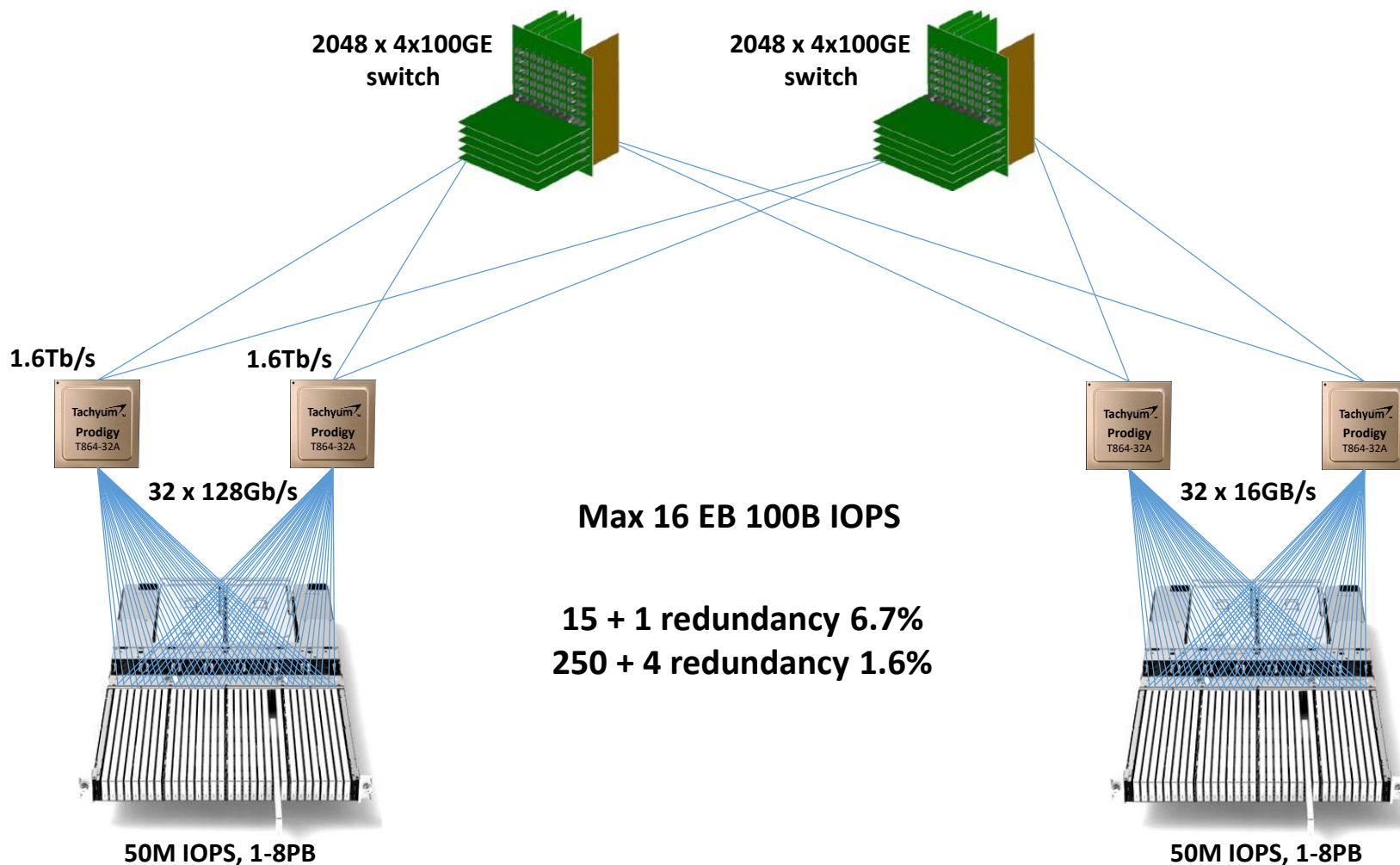


Copper Rack → Fiber Spine

128 x 2x100GE PAM4 switch chip
12U 4K ports x 200GE switch
front-connector-back cards

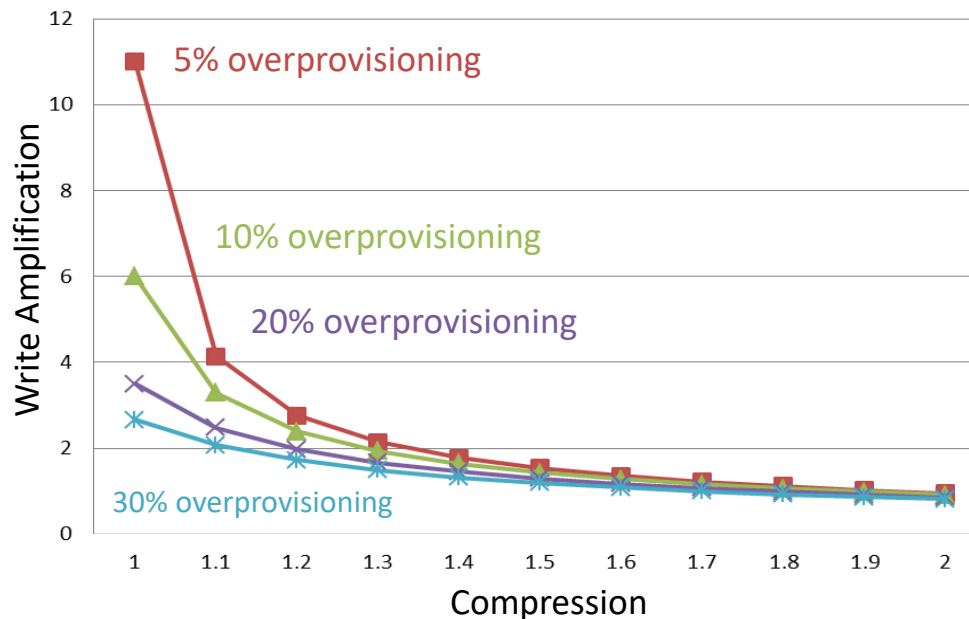


Private Cloud Architecture



10x Effective Life Amplification

- 10x life amplification from compression
 - The compression has non-linear impact on life amplification
 - Example 2:1 compression and 5% overprovisioning giving 10x life amplification
 - SandForce with IBM proved 10x life amplification with 2:1 compression in real life applications
 - Speaker was founder and CTO of SandForce
 - No other SSD controller succeeded in implementing compression based life amplification



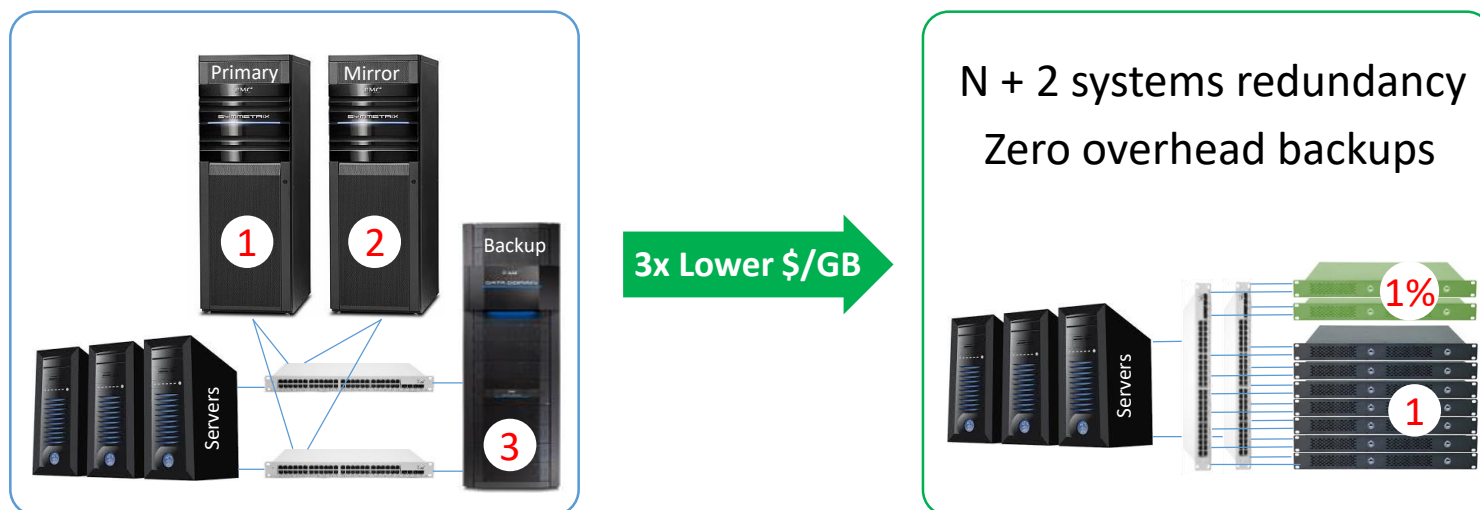
100x Effective Life Amplification

- 2.5x Deduplication and improved compression
 - From 2:1 compression to 5:1 compression and deduplication
 - Invented by Skyera and Pure Storage for primary flash storage
 - Speaker was founder and CEO of Skyera
- 3x One write for protecting against 2 SSD failures instead 3 writes for RAID6
 - It is not compatible with standard SSD use and requires a custom flash controller
 - Garbage collection and compression must be done on system and not SSD level as in SandForce
 - Data are written sequentially; flash of different drives and protections symbols are accumulated
 - Invented by Skyera and Pure Storage for primary flash storage
 - Speaker was founder and CEO of Skyera
- 1.33x Thin provisioning, zero overhead clones and snapshots
 - Invented by Skyera and Pure Storage for primary flash storage
- 100 x life amplification = $2.5 \times 3 \times 1.33 \times 10x$ from compression + recycling



300x Effective Life Amplification

- Typical enterprise storage and private cloud storage uses 3-copy system
 - RAID6 reduces 2-3x performance and by another 2-3x factor during long rebuild times
 - RAID6 does not help when whole rack fails or part of the building get damaged (fire, ...)
 - That is why primary system has mirror system and also backup system
- 3x From system level failure tolerance without need for 3 copies
 - Write data and metadata sequentially across flash in different systems
 - Distributed processing allows for 2-4 complete system failures without data unavailability
 - Tachyon's Prodigy chip has enough spare performance to not show slowdown during rebuilds
 - Processor and network cost is reduced to low enough level that entire solution is cost effective

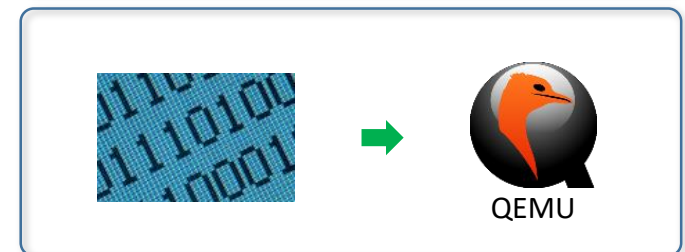
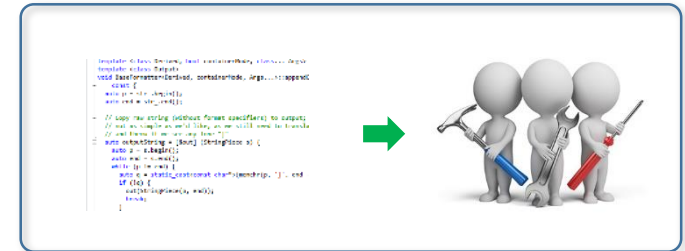


QLC Flash Can Replace HDD in Datacenters

- Assume 300 P/E (Program/Erase) cycles for QLC flash
 - 90,000 effective cycles = 300 x life amplification x 300 P/E cycles
- We need conventional SSD with flash with 90,000 P/E cycles
 - If we place them into existing RAID6 system
 - If we use snapshots, cloned and thick provisioning
 - If we make 3 copies for protecting against system failures
- HAMR disk drives write endurance is limited by laser active lifetime
 - Seagate proved single-head HAMR data writes of over 2PB (20TB drive has 16 heads)
 - So $2\text{PB} * 16 \text{ heads} / 20\text{TB} = 1,600$ full drive writes during lifetime, equivalent to 1,600 P/E cycles
- QLC is lower cost than disk drive in the datacenter with Tachyum chips
 - Disk 11¢/GB: 3 copies 10TB 3.5" \$320 HDD = 9.6¢/GB + 1.4¢/GB system
 - Flash 9¢/GB: DRAMeXchange 32GB USB \$2.5 mCOB = 7.8¢/GB + 1.2¢/GB system
- QLC endurance is sufficient for datacenters with Tachyum chips
 - 300 P/E cycles QLC with Tachyum chips has similar effective endurance as existing conventional datacenter using systems with SSDs with flash endurance 90,000 P/E cycles for typical

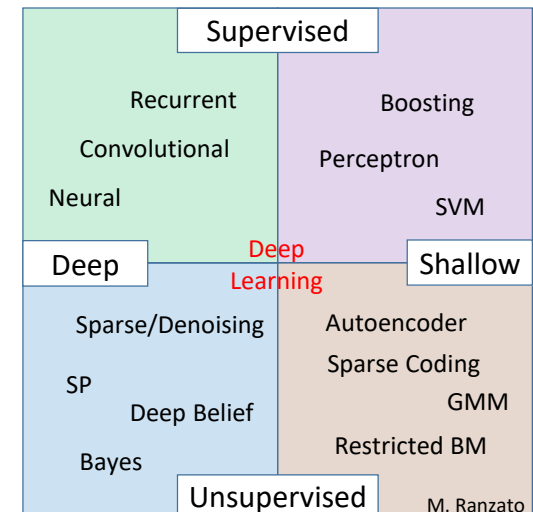
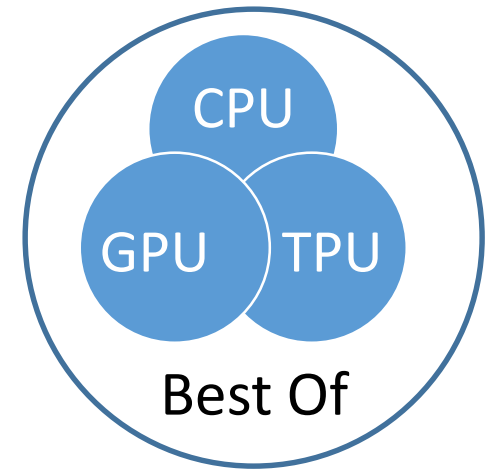
Software Model for Prodigy Chip Customers

- Tachyum does not build systems or software
 - But provides compiler and operating systems
 - Provides IP and libraries to builders of storage systems
 - Provides know-how how to build storage systems
- Tachyum-ported software
 - GCC with Tachyum backend, LLVM in 2019
 - Porting Linux and Free BSD in 2019
 - Device drivers, Boot-loader and Java JIT
- Existing Applications Recompiled
 - Hardware supports strong or relaxed memory ordering
 - Recompiled applications run faster than on Xeon
 - Apache, MySQL, Hadoop, Spark, TensorFlow, ...
- Existing binaries supported via emulators
 - QEMU and emulators transparently launched by Linux
 - Deployment of processor before all applications ported
 - Port CPU intensive application first, other later



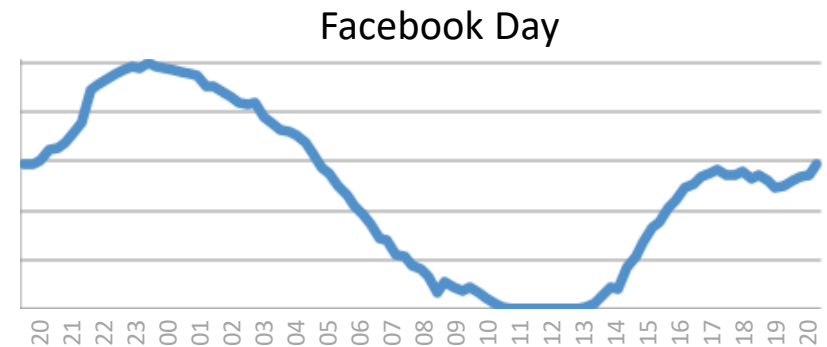
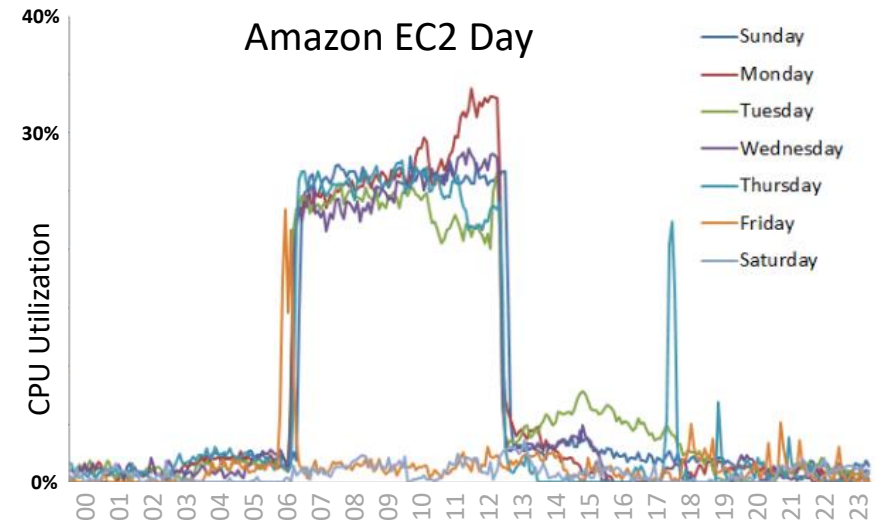
Prodigy: Universal Processor / AI Chip

- Prodigy is a Server/AI/Supercomputer Chip
 - For hyperscale datacenters, HPC and AI markets
- First time humanity can simulate human brain-sized neural networks in real-time
 - Critical for the Human Brain Project
- Prodigy: a Tachyum Architecture
- Outperforms CPU, GPU and TPU
 - CPU: easy to program, costly & power hungry
 - GPU: much faster but very hard to program
 - TPU: faster but more limited apps than GPU



Prodigy: Big AI for Datacenters CAPEX Free

- Universal Processor / AI chip:
10x more AI using idle servers
- Avg. over 24 hours: 60-80% of servers are idle
<5% of servers have AI GPUs
Prodigy enables idle servers to be seamlessly and dynamically reconfigured into HPC/AI systems
- Existing Processors - too slow for AI
therefore, GPU or TPUs are used



Brain Simulation In Hyperscale Datacenter

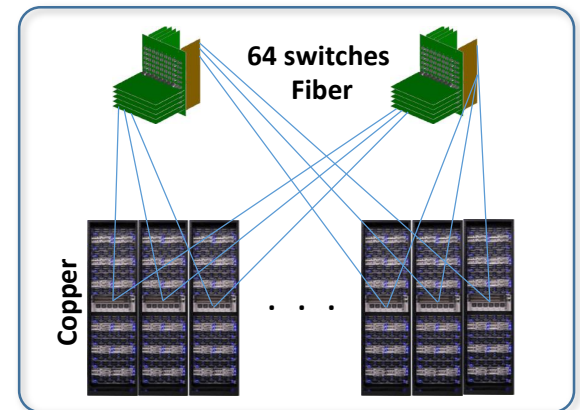
- From Rat Brain to Human Brain real-time simulation

- SpiNNaker system 518,400 processors simulates rat brain
- Human brain simulation requires 1,000x more performance
- The NNSA 20 Pflops Sequoia is 1,542x slower than real-time



- How a system can be built in 2020

- 256K servers, each 4 x 2x100GE with no oversubscription
- Partner's 128 x 2x100GE PAM4 switch chip
- Copper 64 nodes to rack switch, fiber to central switches
- 12U 4K ports x 200GE switch, front-connector-back cards
- Only 1 set of fibers 256 x 2x100 GE vs. 3 to central switches



- 100+ brain-capable datacenters

- Facebook: 100MW datacenter with 442,368 servers
- 40% utilization means 265,420 idle servers
- Use \$100B of underutilized equipment in the world



Prodigy Delivers Low Power Flash Cloud

- Datacenters today consume 2% total electricity

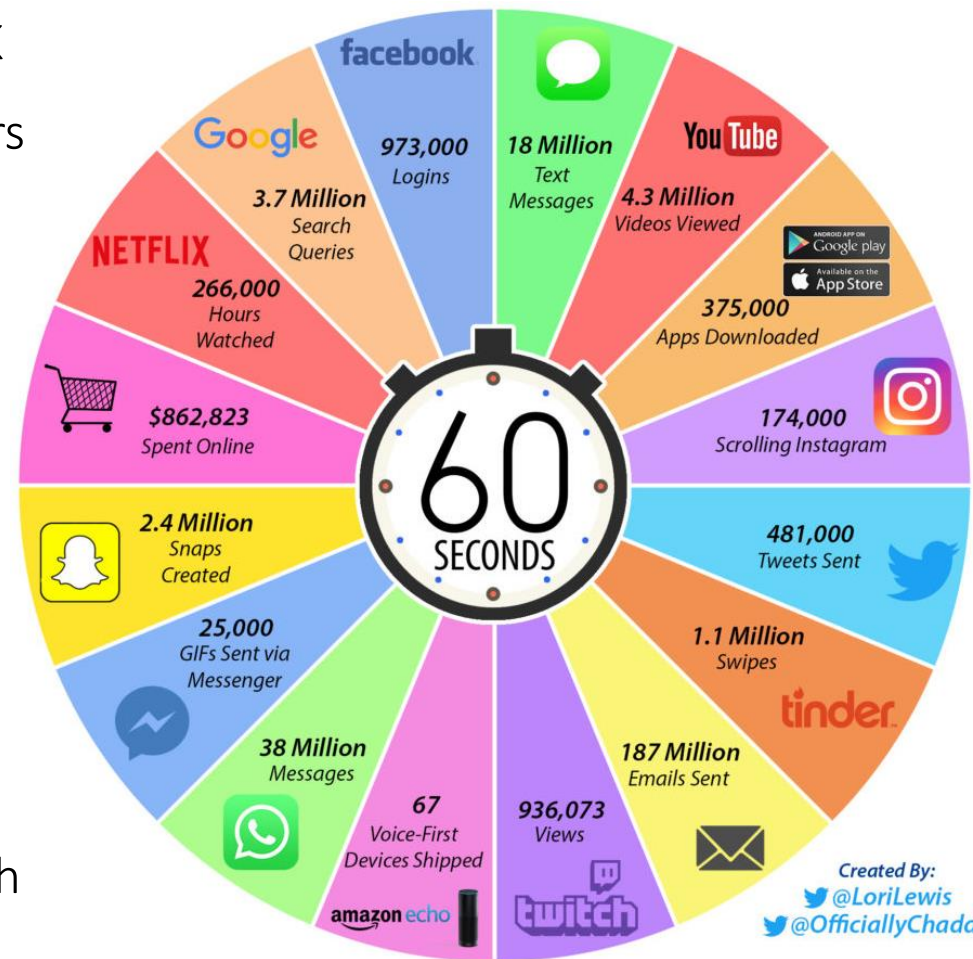
- Consume 40% more power than UK
- Emit more CO2 than world's airliners

- **10% of planet energy by 2030**

- 15% growth: is 2x every 5 years
- 40% of planet energy by 2040

- New Technology is needed

- 10x lower power to continue growth



Tachyum \$10+B Semiconductor Company

Product Faster & 10x more efficient processor than Xeon

Disruption Flash only datacenters below disk drive cost

Status Tape-out 2019, production 2020



Visit us: www.Tachyum.com

Follow us:      



HPC wire silicon startup coming onto the HPC/hyperscale scene with some intriguing and bold claims

451 Research attractive proposition for hyperscale cloud providers, which could potentially build a single architecture that could be repurposed

MICROPROCESSOR *report*
new 64-bit architecture that combines elements of RISC, CISC, and VLIW



\$24B market
10x less power



1st real-time human brain sized neural network sim



Hyperscale/AI/HPC
3x Lower Capex