Liquid Cooling: Exceeding the Limits of Air Cooling to Unlock Greater Potential in High Performance Computing

27 Aug. 2019

Dr. Werner Scholz
XENON Systems, CTO and Head of R&D
werners@xenon.com.au

XENON High Performance Computing

www.xenon.com.au

XENON Systems – Who We Are

Australian company established in 1996.

Global Onsite hardware support and installation services in over 80 countries

Direct Relationship with all major component manufacturers to lower cost and speed up support

- Scientific / Academic Research
- Oil and Gas
- Cloud

- Defence
- Education
- Broadcast

- Finance
- Telecommunication

In-house technical ability to build low volume custom designed servers

Focused on innovation and investment in Research & Development

Subsidiaries:

Mediaproxy
Global leader in compliance logging and transport stream monitoring for broadcast and TV industries.

XDT/Catapult
Software for film and post production industries.

XENOptics
Fibre automation solutions for SDN in data centres

XENON High Performance Computing

XENON Systems – A History of High Performance Solutions

1990 **2000** **2010**

1998
Telstra Digital Video Network
Supplied Video Server equipment

2000
Animal Logic
Render Farm delivered for the "Matrix" movie trilogy

2004
TataSky DTH India
Supplied Video Logging solution

2005
Thales Defence
Supplied Image Generators for the ASLAV simulators

2007
Victorian Partnership for Advance Computing
Supplied HPC Cluster for Advance Computing

2009
CSIRO
(10th in Green500 ranking)
Supplied Australia's first GPU Cluster ("Bragg")

2011
Fox Broadcasting USA
Supplied Video Logging solution

2011
Australian Stock Exchange (ASX)
Supplied Infiniband equipment

2012
Fujitsu
Supplied Infiniband Network for Australia's largest Supercomputer ("Raijin")

2013
Digital Cinema Systems
Upgraded 150 Independent Cinemas in ANZ region

2014
University of Queensland
FlashLite HPC cluster system for high data throughput

2015
Thales Defence
Delivered High Fidelity Solution for Australian Army Tiger Helicopter Simulator

2015
Futuris
XENON drive high performance innovation

2016
Oil Search
Deployed Intel® CAS to speed reservoir modeling

2016
Garvan Institute of Medical Research
NVMe solution that enables high data throughput

2017
The Walter and Eliza Hall Institute
Private Cloud For Next Generation Cancer, Disease & Medical Research

2017
NCI
Supercomputer dedicated to scientific research & technology innovation

2018
Todd Energy (NZ)
HPC Cluster for Reservoir Modelling
Fully managed service

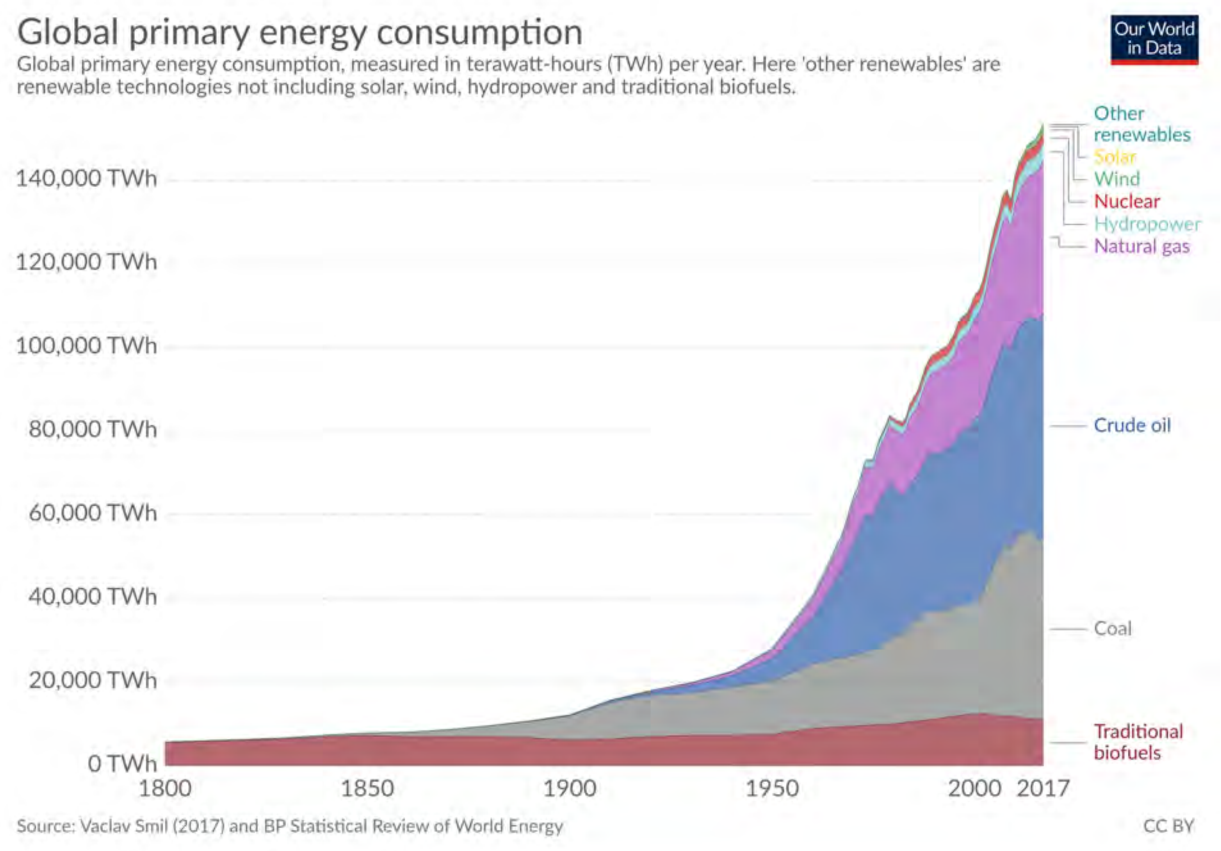
CSL
HPC Cluster based on Cisco UCS

XENON High Performance Computing

Why do we need to drive Data Centre Efficiency?

- Reduce power consumption
- Reduce fossil fuel consumption
- Reduce CO_2 emissions
- Reduce cost
- Save the planet

What is your motivation?



XENON High Performance Computing

Refs:
https://www.theregister.co.uk/2016/07/25/semiconductor_industry_association_international_technology_roadmap_for_semiconductors/
http://www.semiconductors.org/main/2015_international_technology_roadmap_for_semiconductors_itrs/

Why do we need to drive Data Centre Efficiency?

- Reduce power consumption
- Reduce fossil fuel consumption
- Reduce CO_2 emissions
- Reduce cost
- Save the planet

What is your motivation?

Some scenarios predict that computers will consume more power than the world can produce by 2040.
(Semiconductor Industry Association Roadmap 2015)

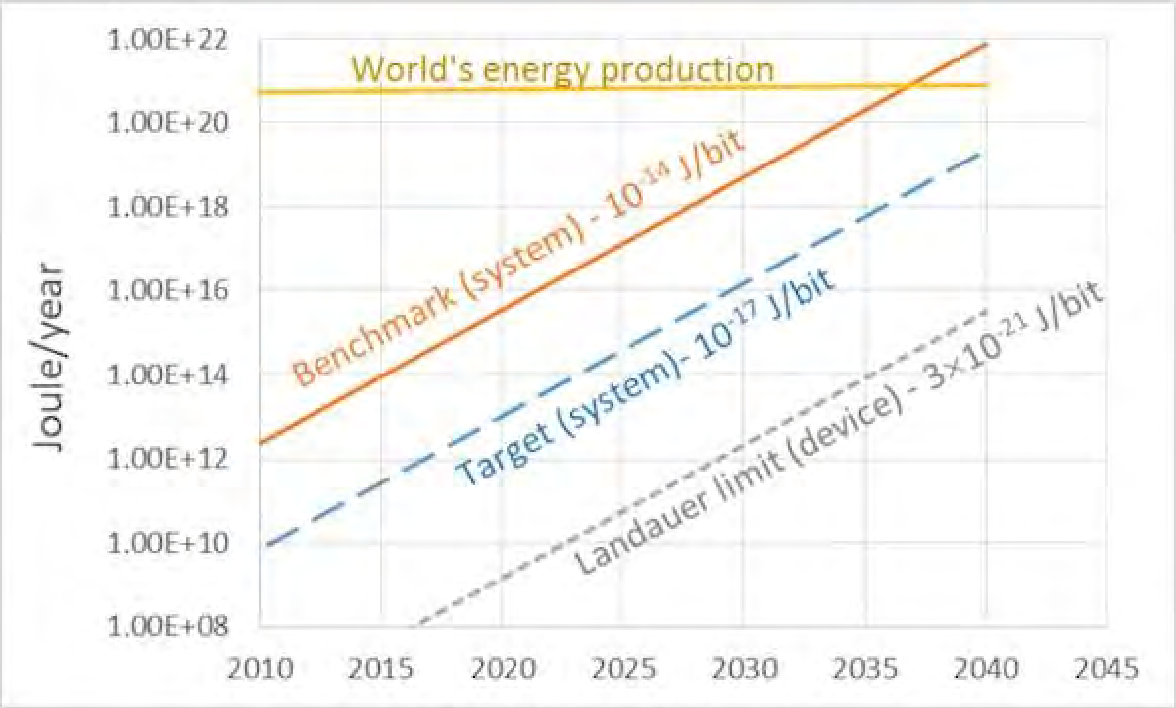
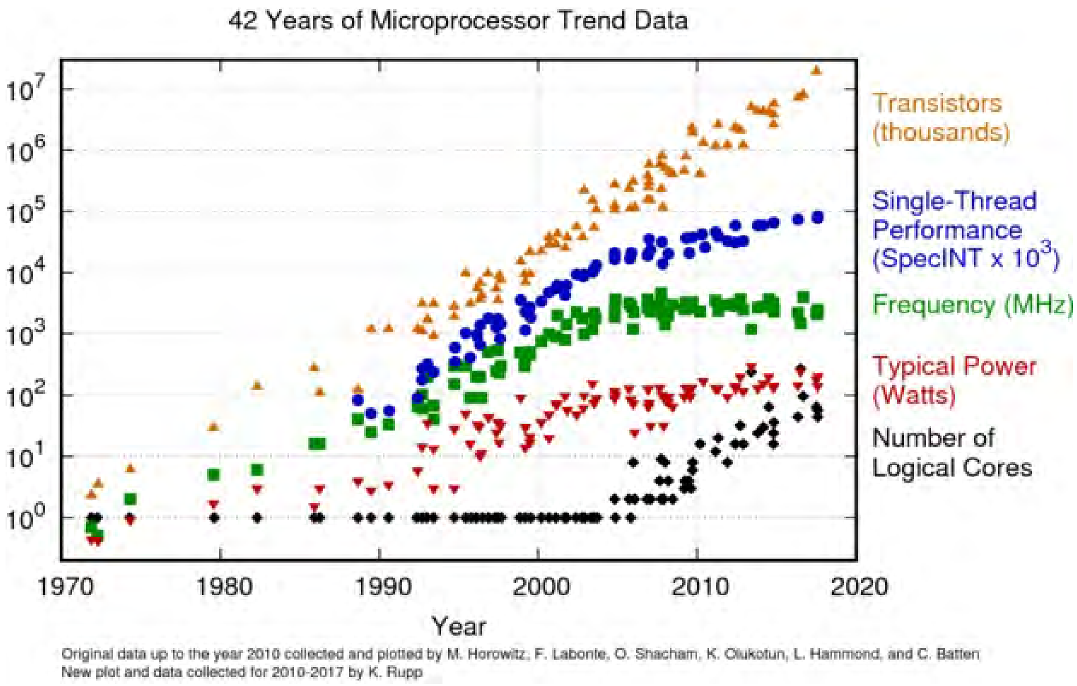


Fig. A8. Total energy of computing.

XENON High Performance Computing

Refs:
https://www.theregister.co.uk/2016/07/25/semiconductor_industry_association_international_technology_roadmap_for_semiconductors/
http://www.semiconductors.org/main/2015_international_technology_roadmap_for_semiconductors_itrs/

Component Trends and Demand are Challenging

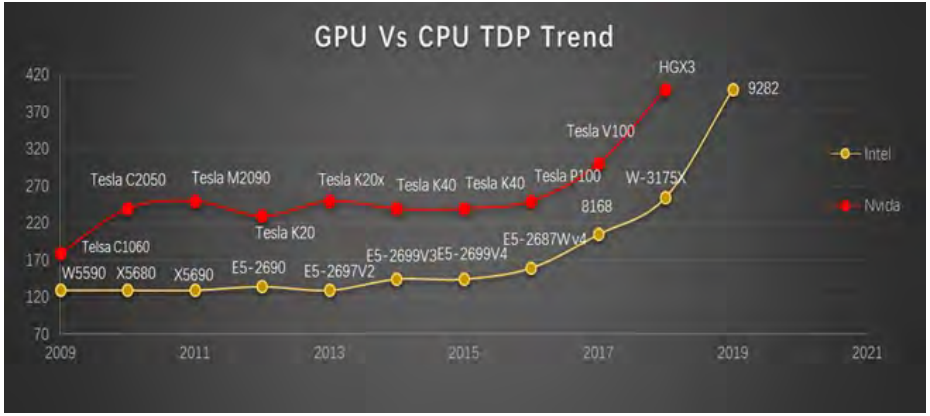


- Single thread performance increases are slowing
- Number of logical cores and transistors per socket is increasing
- Power efficiency gains per core are slowing
- Power per socket is increasing

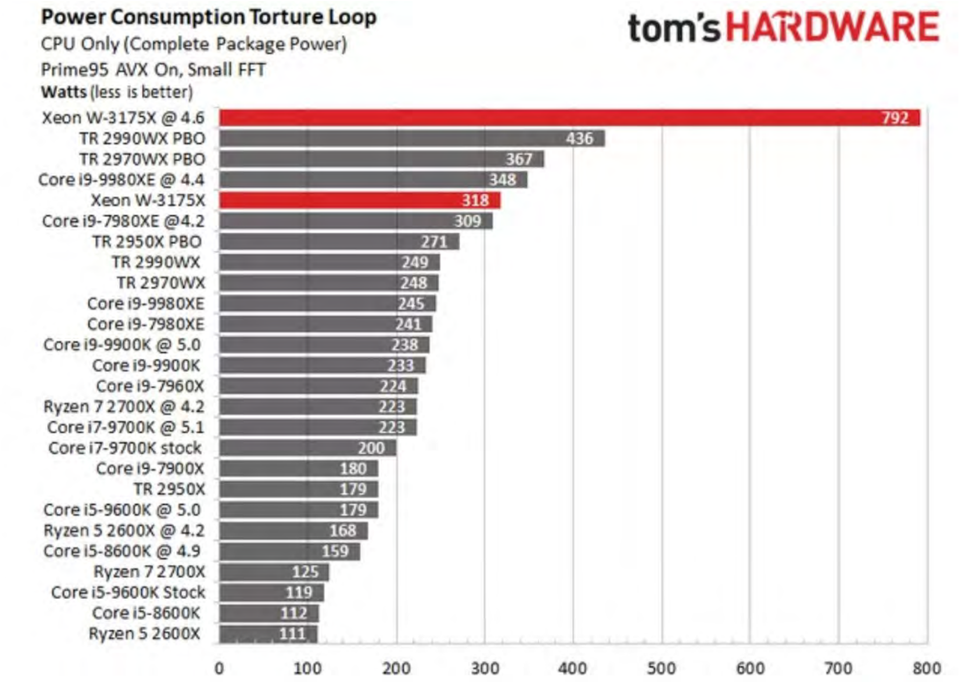
XENON High Performance Computing

Refs:
https://www.karlrupp.net/2018/02/42-years-of-microprocessor-trend-data/

Component Trends and Demand are Challenging



- GPU: currently up to 300 W per device
- CPU:
  - Jumping from typ. 150 W to
  - 225 W (AMD EPYC 7742 with 64 cores)
  - 255 W (Intel® Xeon® W-3175X) and recently
  - 400 W (Intel Xeon 9282 “Cascade Lake AP” with 56 cores)
  - even higher for overclocked CPUs



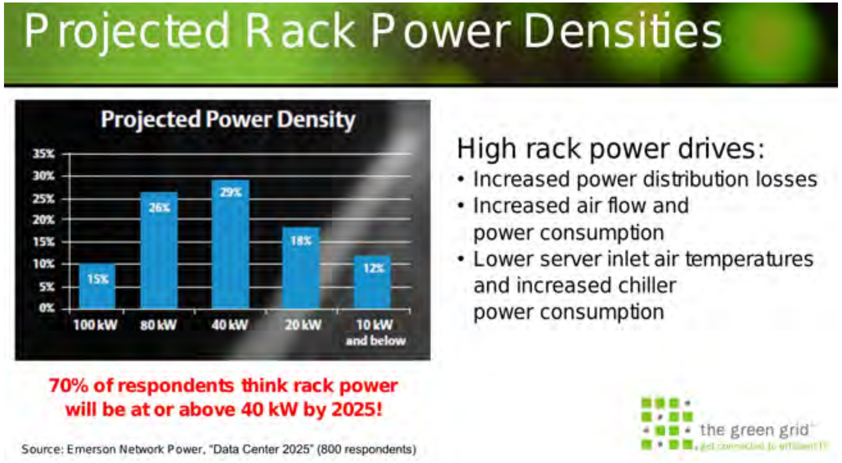
XENON High Performance Computing

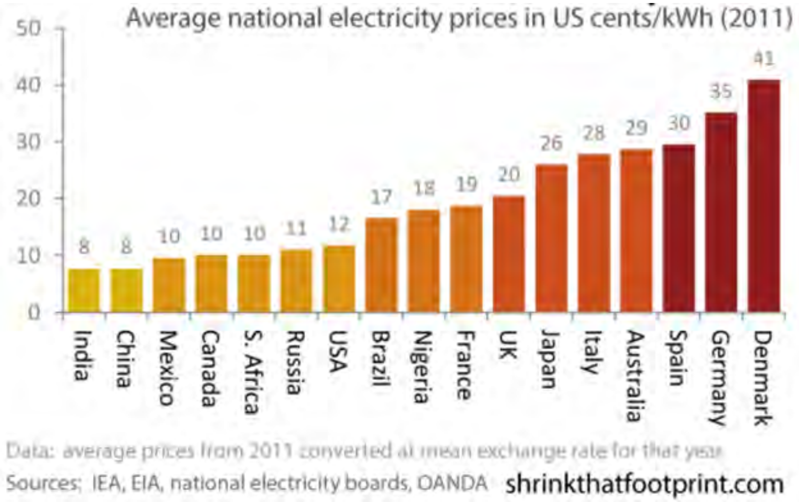
Refs:
https://www.datacenterdynamics.com/opinions/its-not-just-about-chip-density-considering-liquid-cooling-for-your-data-center/
https://www.tomshardware.com/reviews/intel-xeon-w-3175x-cpu,5976-3.html

Impact on Data Centres

Challenges:

- Power density/power per rack increases
- Power feeds
- Power distribution
- Power cost
- Hot spots
- DC/Rack/Component Cooling





XENON High Performance Computing

Refs:
11 PUE Data Center for the Masses A Webcast by The Green Grid.pdf
https://www.thegreengrid.org/

Setting a Target: PUE = 1

PUE = Power Usage Effectiveness

DCIE = Data Center Infrastructure Efficiency

$$PUE = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}} = 1 + \frac{\text{Non IT Facility Energy}}{\text{IT Equipment Energy}}$$

DCIE = 1/PUE

Data Centre	PUE
Google (2008)	1.21
Microsoft (2008)	1.22
Google (2010)	1.16
Facebook (2011)	1.08
Typical Data Centre (2011)	1.50
Microsoft (2012)	1.20
Google (2012)	1.14
Switch Supernap 7 (2015)	1.18
Facebook (Prineville 2015)	1.08
Allied Control (Bitfury 2015) (2PIC)	1.02
Green IT Cube (2016) (RDHX)	1.07
Supermicro (2017)	1.06
Facebook Prineville DC (yesterday*)	1.04
Facebook Forest City DC (yesterday*)	1.08
Facebook Lulea DC (yesterday*)	1.16
Facebook Altoona DC (yesterday*)	1.14

Refs:
https://gigaom.com/2012/03/26/whose-data-centers-are-more-efficient-facebooks-or-googles/
https://googleblog.blogspot.com/2012/03/measuring-to-improve-comprehensive-real.html
https://www.datacenterknowledge.com/archives/2013/04/18/facebook-unveils-live-dashboard-on-pue-water-use
https://www.google.com/about/datacenters/efficiency/internal/index.html#measuring-efficiency
https://www.facebook.com/notes/facebook-engineering/designing-a-very-efficient-data-center/10150148003778920/
https://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf

*) Live Facebook data centre dashboards
https://www.facebook.com/PrinevilleDataCenter/app/399244020173259/
https://www.facebook.com/ForestCityDataCenter/app/288655784601722/
https://www.facebook.com/LuleaDataCenter/app/115276998849912/
https://www.facebook.com/AltoonaDataCenter/app/602730866540556/

How to Drive Data Centre Efficiency Gains?

Some ideas:

- Improve Data Centre design
- Location selection (more on that later)
- Improve component selection (for efficiency)
- Improve component design (offload, BIGLittle)
- Develop smarter algorithms:
  Do more with less (but doing more with more is becoming cheaper faster...)
- Smarter load distribution: In-Server, In-Rack, In-DC (avoid hot spots)
- **Improve Cooling Efficiency**

XENON High Performance Computing

Refs:
https://www.asetek.com/data-center/data-center-oem-partners/intel/
https://www.xenon.com.au/solutions/high-frequency-trading-solutions/

Liquid Cooling Options for Every Situation

Liquid cooling includes a whole family of options

In-server

- Closed Loop Liquid Cooling
  - Actively pumped
  - Passive solutions
- Heat Pipe and Solid Conduction

In-rack

- LAAC (Liquid Assisted Air Cooling)
- Rear Door Heat Exchangers
- In-rack CDU

In-Data Centre

- Immersion Cooling
  - Single phase immersion
  - 2-phase immersion
  - “Full DC immersion”



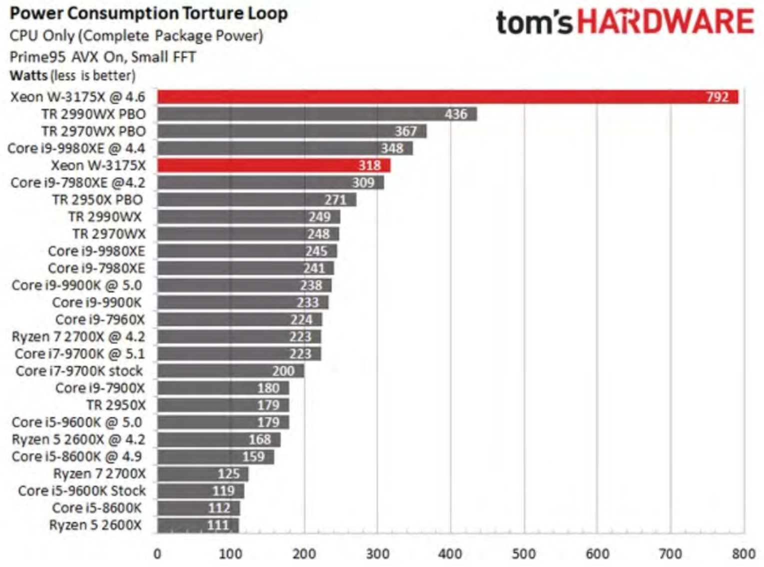


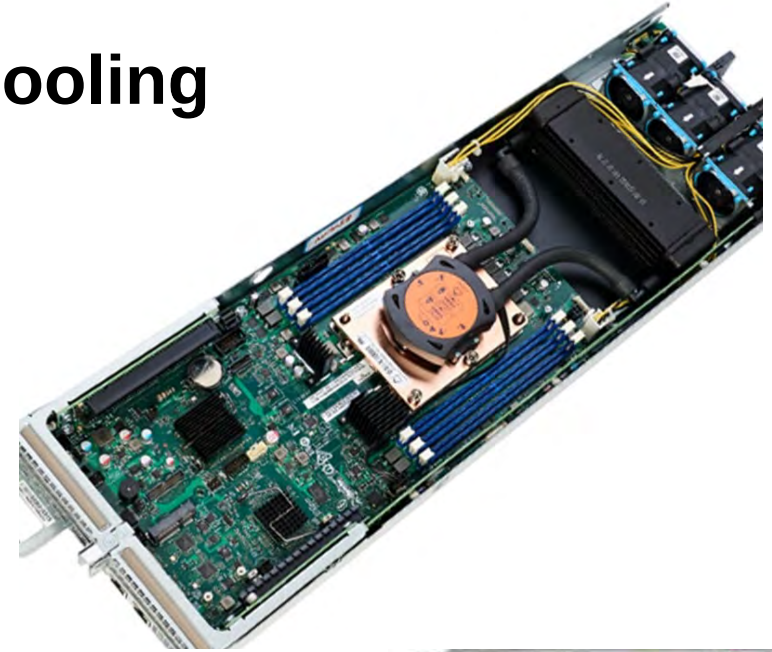

XENON High Performance Computing

In-server Closed Loop Liquid Cooling

Direct-to-chip liquid cooling enables

- Higher cooling efficiency
- Higher power CPUs/GPUs
- Higher density
- Higher performance (turbo/overclocking)





Specifications

MODEL	XENON RADON Solo 1U RX991i 4Cores 5.2GHz HFT SERVER	XENON RADON Solo 1U RX993i 10Cores 5.0GHz HFT SERVER	XENON RADON Solo 1U RX998i 18Cores 4.8GHz HFT SERVER
PERFORMANCE	Overclocked up to 5.2GHz with 4Cores Active	Overclocked up to 5.0GHz with 10Cores Active	Overclocked up to 4.8GHz with 18Cores Active
PROCESSOR BASE FREQUENCY (CPU)	4.3GHz	3.3GHz	2.6GHz
HYPER-SPEED FREQUENCY (CPU)	5.2GHz	5.0GHz	4.8GHz
MEMORY BASE FREQUENCY (RAM)	2666MHz	2666MHz	2666MHz
HYPER-SPEED FREQUENCY (RAM)	3200MHz	3200MHz	3200MHz
CPU	Single Intel® Core™ i7-7740X (Kaby Lake-X) 4Cores 8MB Cache 112W	Single Intel® Core™ i9-7900X (Skylake-X) 10Cores 13.75MB L3 Cache 140W	Single Intel® Core™ i9-7980XE (Skylake-X) 18Cores 24.75MB Cache 165W

Refs:
https://www.tomshardware.com/reviews/intel-xeon-w-3175x-cpu,5976-3.html
https://www.asetek.com/data-center/data-center-oem-partners/intel/
https://www.xenon.com.au/solutions/high-frequency-trading-solutions/

XENON High Performance Computing

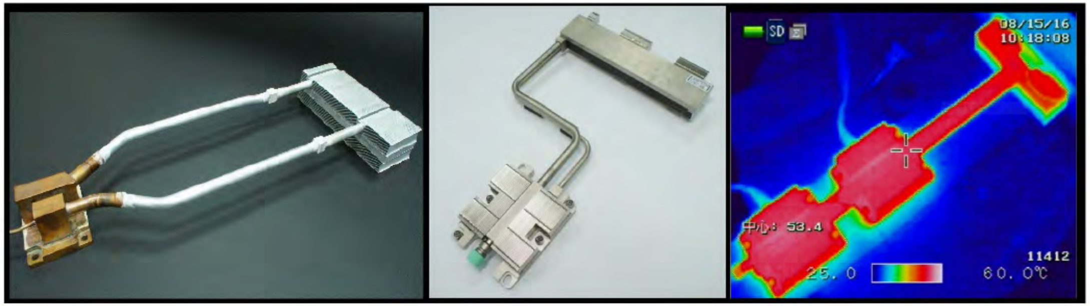
Passive In-Server Cooling

1) Passive closed loop liquid cooling
   - Cover longer distances (to condenser)
   - Larger condenser for better efficiency
   - Enables higher TDP
2) Heat Pipe and Solid Conduction

No active parts

No risk of failure of active parts

Fully sealed systems: no leakage



Refs
https://www.calyos-tm.com/processor-cooling-for-datacenter :
http://www.lemtech.com/products/heatsinks/#Process

In-Rack Cooling

Rear Door Heat Exchangers

- “Hot aisle” confinement within rack
- Eliminates hotspots and “overcooling”
- Retrofit onto existing racks usually possible
- Easy installation
- “Zero U”
- Small space requirement (additional rack depth)

InRack LAAC (“Liquid Assisted Air Cooling”)

- Rack-mounted 2U cabinet containing liquid-to-air (L2A) heat exchanger
- Capable of rejecting up to 6.4kW of total processor power
- Captures 60% to 80% of server heat with Asetek D2C cooling loops

InRackCDU, OnRackCDU

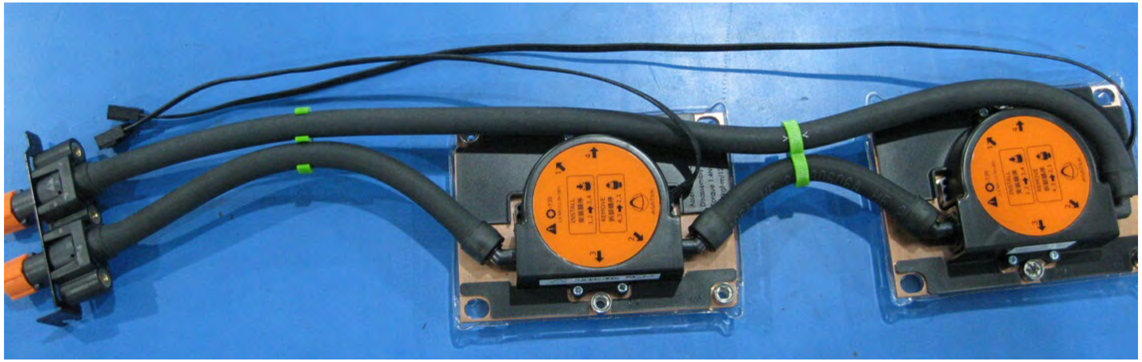
- Liquid-to-liquid (L2L) heat exchanger
- Rejects up to 80kW of heat from the rack
- Captures 60% to 80% of server heat with Asetek D2C cooling loops
- 2.5x-5x increases in rack density

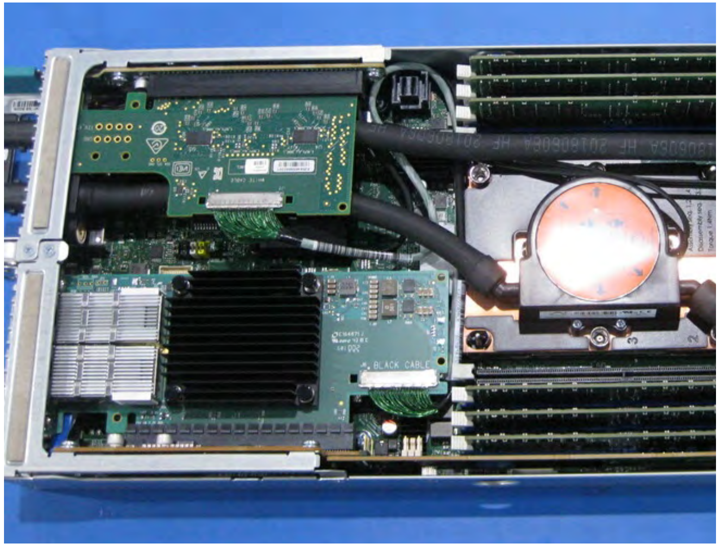
Refs:
https://www.asetek.com/data-center/technology-for-data-centers/inracklaac/

Recent Project

OpenStack HPC Cluster

- Intel 0.5U server (2U 4N)
- Intel 2U servers for Ceph NVMe Storage
- Supermicro 1U server with 4x NVIDIA V100
- Water cooled CPUs and GPUs





	Air cooled	Liquid cooled
CPU LINPACK		
CPU1 temp	73 C	52 C
CPU2 temp	76 C	55 C
GPU LINPACK		
Performance (TFLOPS)	18.44	19.07
GPU power (typ.)	300 W	300 W
GPU power (max.)	375 W	379 W
GPU temp. (typ.)	50-60 C	40 C
GPU temp. (max.)	65 C	49 C

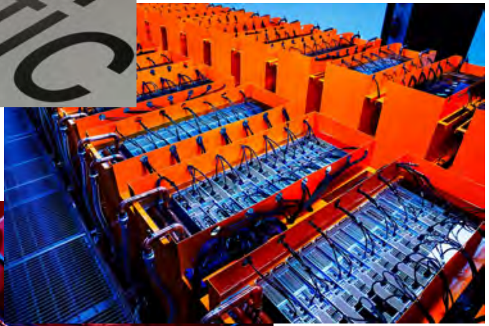
Immersion Cooling

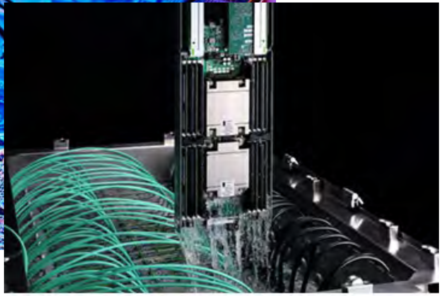
1) Single phase immersion

- Reduced number/elimination of fans
- Data Centre temp. and humidity can increase (to comfortable levels)
- Data Centre noise level reduction
- Power reduction
- Cost reduction
- Reliability increases
- Fewer/no fans / fan failures
- Immersion cooling reduces corrosion risk, reduces failures, increases component life

2) 2-phase immersion

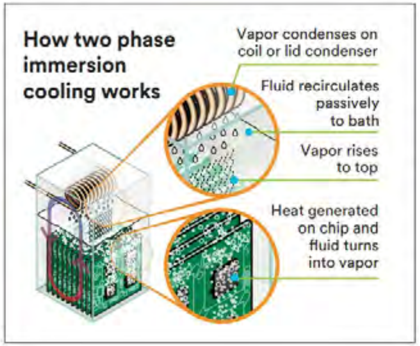
- Same as above
- Hermetically sealed enclosure
- Challenges with vapor management, servicing

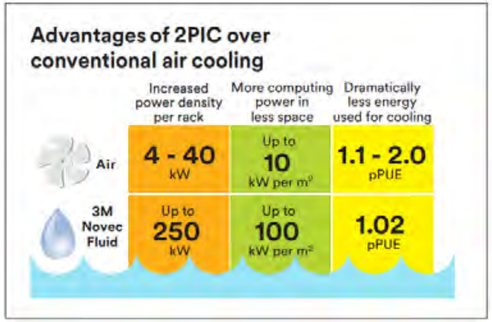




Allied Control 2PIC Projects

	Generation 1	Generation 2	Generation 3
Date	November 2012	October 2013	October 2015
Size/Design	70 kW	500kW 25kW racks scalable to 225kW per rack	40+ MW comprising 250 kW flat-lying tanks
Hardware	FPGA	ASIC compute clusters	ASIC compute clusters
Fluid	3M Novec 7100	3M Novec 7100	3M Novec 7100
PUE	1.02	1.02	1.02
Compute power per ft²	.25 kW/ft²	1.01 kW/ft²	3.23 kW/ft²





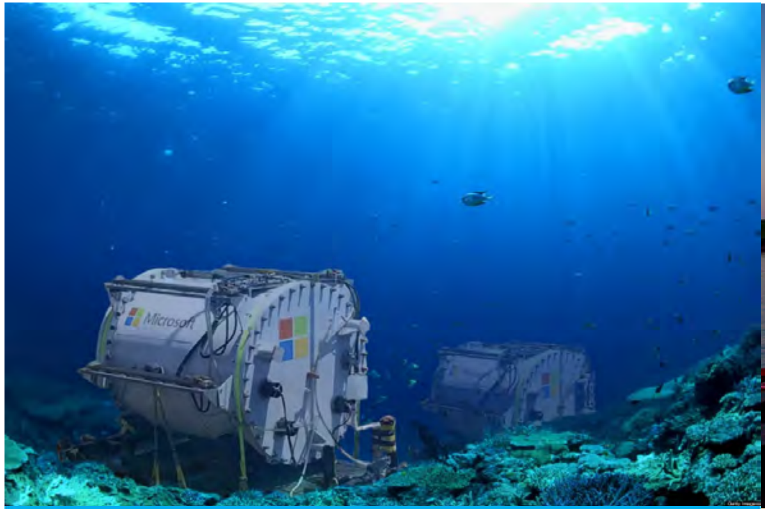
Refs:
http://vsc.ac.at/presscorner/download-area-for-vsc-publications-logos-photos/#c495945
https://dug.com/dug-cool/
https://www.fujitsu.com/global/about/resources/news/press-releases/2018/1114-01.html
http://multimedia.3m.com/mws/media/1127920O/2-phase-immersion-coolinga-revolution-in-data-center-efficiency.pdf

“DC Immersion Cooling”

Microsoft Research

Project “Natick”

- New DC deployment in 90 days
- Deploy close to urban centers (on coast)
- Free cooling
- Servicing challenges?




Refs:
https://natick.research.microsoft.com

XENON
High Performance Computing

How to Further Drive DC Efficiency Gains

- Measure, measure, measure (power, PUE, load, temp., humidity)
- Improve Thresholds and Sensitivities
  - Increase DC temp. as much as possible
  - Increase humidity range as much as possible
- Improve DC design and airflow management
  - Optimize tile layout and cold air outlets to match IT load
  - Blanking plates and side panels
  - Barriers
  - CRAC unit return air inlet design
  - Hot and Cold aisle containment
  - Use free cooling (as climate permits)
  - Invest in CFD modelling (ROI could be <6 months)
  - Sensor lights to reduce light load
  - Reduce power conversion
  - Use high efficiency UPS

From a Google Case Study

PUE Improvements for 5 Network POPs

PUE Improvement	POP 1	POP 2	POP 3	POP 4	POP 5
Starting point	2.4	2.2	2.2	2.4	2.4
Implementing "immediate improvements"	2	2	2	2.2	2.2
Tightening cold aisle	1.9	1.8	1.8	2	2
Adding RA plenum to CRAC & 2 unit operation	1.7	1.7	1.7	1.7	1.7
Adding new controller	TBD	1.5	TBD	1.6	1.5

- UPS power 80kW at beginning and 86kW at end
- PUE data collected every second at power quality meters; daily PUE value averages 86,400 data points
- Awaiting data in some locations

Retrofit result: PUE decreased from 2.4 to 1.5

Google

PUE Improvements and ROI Summary

US East Coast Data Center	PUE	Capital Investment [$USD]	PUE Improvement	Savings / Month [$USD]	ROI [Months]
Starting point	2.4	N/A	N/A	N/A	N/A
Implementing "immediate improvements"	2.2	–	0.2	$1,238	–
Tightening cold aisle	2	$12,000	0.2	$1,238	9.7
Adding RA plenum to CRAC & 2 unit operation	1.7	$5,000	0.3	$1,858	2.7
Adding new controller	1.5	$8,000	0.2	$1,238	6.5

Recommend thorough analysis of your site before implementing.

- Cost does not include company manhours for work (i.e. installation of blanking plates) and CFD analysis.
- Simple payback ROI based on $0.10 per kw-hr.

Return: $67k/year in savings on $25k retrofit

Google

Refs:
Joe Kava, "Central Network Room (CNR) PU Efficiency Project - A True Story."
2011 Data Centre Efficiency Summit held on May 24 in Zürich, Switzerland.
https://www.youtube.com/watch?v=APynRrGuZJA
http://static.googleusercontent.com/media/www.google.com/en/us/corporate/datacenter/dc-best-practices-google.pdf

Conclusions

Benefits of Liquid Cooling

- Reduced number/elimination of fans and other air cooling components
- Data Centre temp. and humidity can increase (to comfortable levels)
- Data Centre noise level reduction
- Power reduction
- Cost reduction
- Higher power density systems and racks
- Performance increases (turbo/overclocking)
- Reliability increases
  - Fewer/no fans / fan failures
  - Immersion cooling reduces corrosion risk, reduces failures, increases component life

Integrated with careful Data Centre Design further gains can be made

Contacting XENON Systems

XENON Systems Pty Ltd
10 Westall Road
Clayton South, VIC 3169
P   +61 3 9549 1111
F   +61 3 9549 1199
info@xenon.com.au
www.xenon.com.au
ABN: 39 074 339 316

Contact Details
Werner Scholz
CTO and Head of R&D
03 9549 1140
werners@xenon.com.au

Michael Shing
Business Development Manager
03 9549 1144
michaels@xenon.com.au

Thank you!

Questions?

Dr. Werner Scholz
XENON Systems, CTO and Head of R&D
werners@xenon.com.au

XENON High Performance Computing
www.xenon.com.au