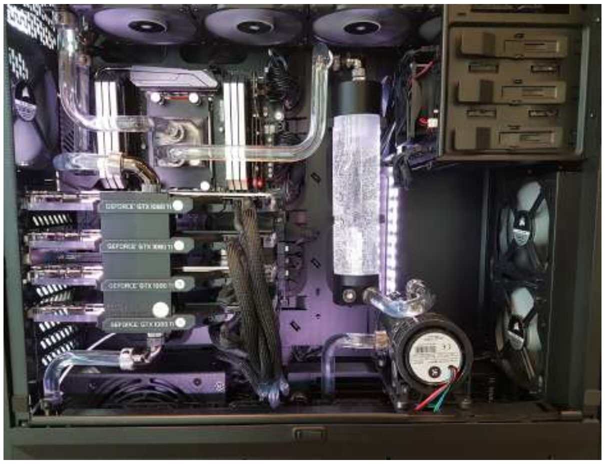


Comprehensive GPU workstation roundup - Late 2017

Shintaro Aibara, Dari Kimanius, Bjorn Forsberg, Stefan Fleischmann

Entry Level Content
Enthusiast Level Content

Dear Friends and Colleagues,

Mid-2016, we released a page in our blog pointing towards a couple of potential builds suitable for cryo-EM processing. With the ever updating hardware, and in particular in 2017 with the release of ever more powerful CPUs and GPUs. We have decided to bring to you a more updated recommendation of workstations suitable for processing in RELION and other structural biology software.

In this report, we have compared a number of systems, including our previous “cost-efficient” workstation (http://www.cryoem.se/hardware-for-a-low-end-workstation/) and newer AMD-based workstations that offer compelling solutions for multi-GPU systems. We have extensively investigated the effects of SSDs (both SATA and NVMe), graphics cards (980 Ti vs 1070 vs 1080 Ti), and CPU on overall runtimes, using 3D classification and auto-refine with the plasmodium ribosome benchmark dataset.

In a nutshell, the following conclusions can be drawn:

- Broadly speaking, GPU accelerated RELION is still largely CPU bound and so the graphics card choice is not as important as one may have thought previously.

- With a high-end graphics card like the 1080 Ti, you reap the most benefit from having a high-end CPU such as Threadripper 1950X

- It is important to have a good SSD as a scratch space, as this is the single biggest improvement one can get

Recommendation ver. Late 2017 - Air cooled

CPU: AMD Threadripper 1950X or 1920X
CPU Cooler: Noctua U12S TR4 SP3
Motherboard: ASUS Zenith Extreme or Asrock Fatal1ty X399
RAM: 96 GB or 128 GB of DDR4 (ideally a kit supported by the motherboard QVL)
PSU: Corsair AX1500i or EVGA Supernova 1600W
Case: Corsair Air 740
Storage: 2 SSDs (one for the OS, one for scratch) - SATA or NVMe
Graphics card: Any 4x or 2x rear-exhaust graphics card (if buying, 1080 Ti)

CPU Choice

The AMD Ryzen Threadripper platform (X399 Chipset) offers a competitive alternative to Intel's new Skylake-X platform (X299 Chipset). These are both high-end CPU platforms that offer multiple CPU cores and PCIe lanes that support multi-GPU setups. It is notable that for the $1000 USD price point, one could buy a 16-core AMD Ryzen Threadipper 1950X or an 12-core Intel Skylake-X 7900X. While Intel offers an 18-core option, the 7980XE, it costs $700 USD more than the TR1950X and we were not able to test this CPU.

The TR1950X and TR1920X have 60 PCIe lanes available for onboard devices. In this configuration, graphics card can operate in a genuine 16x/8x/16x/8x (48 lanes) leaving 12 more lanes for up to three M.2 NVMe SSD drives (4 lanes each).

Most X399 boards allow for up to four 2-slot graphics card to be installed, whereas one has to be more selective with X299 boards since some boards do not allow you to physically install four dual-slot cards. Furthermore, even the higher-end Skylake-X CPUs only have 44 PCIe lanes, meaning they typically operate at 16x/8x/8x/8x mode. While there are some boards supporting 4-way graphics cards, most do not have dedicated lanes for NVMe drives, which instead go through the chipset DMI connection.

We have assembled two systems using AMD's Threadripper since we anticipate that these processors offer the best price-to-performance for a quad-graphics card system. We have used two boards, the Asrock X399 gaming fatal1ty and the ASUS Zenith extreme.

Graphics card choice

While open-shroud “internal exhaust” graphics card tend to be equipped with beefier heat-sinks and larger fans, these cards will exhaust their heat into the chassis and may perform better in a single or perhaps dual configuration, but would definitely be choked for air in a quad graphics card system.

The GeForce GTX 1080 Ti is the currently the highest end consumer-grade NVIDIA card on the market. Equipped with 11 GB of onboard VRAM memory, offers in our opinion, the best performance.

The onboard memory of the graphics card ultimately dictates the maximum box size that can be refined using RELION, and the 11 GB of the GTX 1080 Ti will allow approximately for a box size of ~500 pixels. With this in mind, the 1070, 1080 and the more recently announced 1070 Ti is in the same “weight class” considering they all have 8 GB of GDDR5 memory allowing for a maximum box size of ~450 pixels.

The only consumer-grade graphics card to exceed the specifications of the GTX1080Ti is the Titan Xp, which boasts 12 GB of onboard memory, however given the steep increase in cost, we do not recommend these cards for use with RELION. Similar comments can be put towards professional-grades cards such as the Tesla series.

While we do think that the 1080 Ti provides fantastic performance, mid-tier 10-series cards such as the 1070 and previous generation cards such as the 980 Ti also perform with compelling numbers. So unless there is a need for more onboard VRAM, one should not feel the need to upgrade the GPU with every new release.

While NVIDIA have announced they do not intend to release a the next GPU architecture on the consumer-end for a while, we are hopeful that they will be released sometime in 2019. How these cards perform can only be wait to be seen.

[Odin, Thor, Loki, Sif, Idun, Freya] “Cost-efficient” workstations refreshed

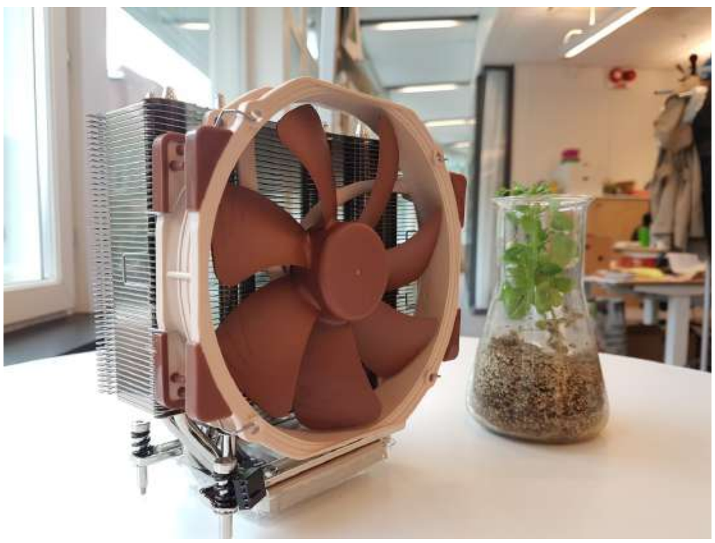
Although hardware is refreshed almost every six months, one should keep in mind that it is not necessary or recommended to update all systems on every release. We have decided to keep our Intel 6900K systems, which starred in the previous “cost-efficient” workstation guide. We have however, updated a few aspects to accommodate our needs

Previously, we had a single SATA SSD on which the operating system was installed and also the scratch space to keep particles (see below). However, we encountered situations where if the scratch space filled, the whole system would seize up. So we have added an NVMe M.2 SSD to have a dedicated scratch space that is not on the same partition as the operating system.

Since we typically work with fairly large boxes, and have found the onboard memory of the 980Ti’s to be left somewhat to be desired; and so we have upgraded 3 of our 6 stations to have dual 1080Tis and 3 to have 1070s.

NOTE We have found that with two GTX1080Ti cards, our 700W power supply from CoolerMaster cannot provide enough power for a 100% stable system, and would reproducibly shut down during one of our benchmarks. Upgrading the power supply to a 850 W power supply has resolved this, and so we would recommend getting a 750~850W power supply for a dual-1080Ti system.

[Zephyrus] AMD Ryzen Threadripper 1920X System - “GPU sink” workstation



Since our “Nordic god” workstations received graphics card upgrades, we have found an excess of 980 Ti’s with no home. We have decided to utilize a cheaper 12-core AMD Ryzen Threadripper 1920X (TR1920X) to bring these excess graphics cards to good use.

This CPU was paired with an ASROCK X399 Gaming Fatal1ty board. This board is particularly desirable for our needs since it has native 10-gigabit ethernet connectivity, meaning that if you have a capable network infrastructure, this would give unrivalled options for high-bandwidth transfer of data.

We have opted for a fairly beefy CPU cooler, the Noctua U14-S TR4 SP3, a cooler specifically designed for the Threadripper CPU and is comparable to even high-end closed-liquid-coolers in terms of cooling performance. It is important to note that some boards such as the ASUS Zenith Extreme have the first PCIe slot too close to the CPU socket and therefore the U14-S would not be suitable. If you are concerned, then we would recommend the Noctua U12 TR4 SP3 Instead. **Please note: There are Noctua coolers with the same name but not the “TR4 SP3” variant, these are not suitable for Threadripper.**

To house all of the above, we have chosen a Corsair Air 740 case, an updated version of the previous high-end workstation. We have equipped our case with three intake fans and 3 exhaust fans. To balance the air pressure within the case, we have used a low-noise adaptor on the exhaust fans to lower the rpm of the exhaust to ensure we have more incoming air than exhaust even when the graphics cards are engaged.

[Poseidon] AMD Ryzen Threadripper 1950X System - “High-end” workstation

The TR1950X is the flagship CPU from AMD this year. We have incorporated this into our high-end workstation, equipped with four GTX1080Tis and three NVMe M.2 SSDs (one for the operating system, and two in RAID0 as scratch space).

We have used the ASUS Zenith Extreme motherboard for this build. In particular we are particularly fond of the DIMM.2 slot, a feature specific to the ASUS boards that allow accessibility to 2 of the M.2 drives even if four graphics cards are installed. This would mean that even in an event where there is a failure of the SSD, the drives can be replaced without unmounting the graphics card (which we hypothesize to be the most likely component to fail in this system - thankfully none have failed as of date).

However, we have chosen this motherboard, case, and name with ulterior motives, so we would not necessarily recommend this case/motherboard choice if you are just after a simple air-cooled system (See Appendix I). Read more below to find out why we chose this behemoth of a case for this system.

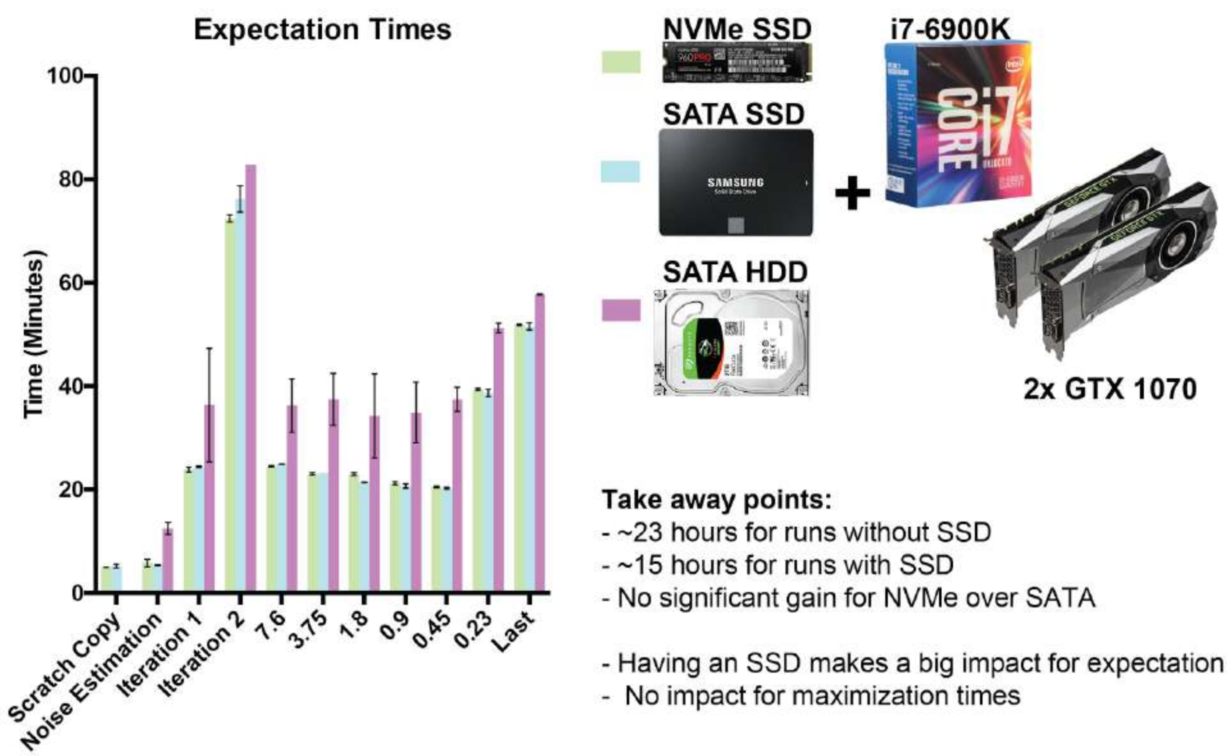
A comment about system memory

One consideration is that with a four graphics card setup, one requires more system memory than the previous recommended 64 GB. Ideally it would be desirable to fully fill the system with 128 GB of memory, but we have experienced difficulties getting 128 GB of memory to be detected by the operating system, or getting the system to turn on at all. We found that 96 GB to be sufficient for most cases, and would recommend this as a starter.

If possible, we would recommend you to find a 128 GB memory kit that is listed on the QVL support memory list for the particular X399 board you have chosen. However, they are fairly rare to come-by. Currently we are making do with 96 GB of system memory (6 x 16 GB sticks).

A SSD is a very worthwhile investment
In our group, we typically store our raw data and RELION project directories on slow mechanical hard drives. While these are relatively cost effective solutions (good price per gigabyte, and also consume no power when unplugged) suffer from the fact that their read/write performance is limited by the mechanical head. Our drives cap at a read/write of 200~250 MB/sec.

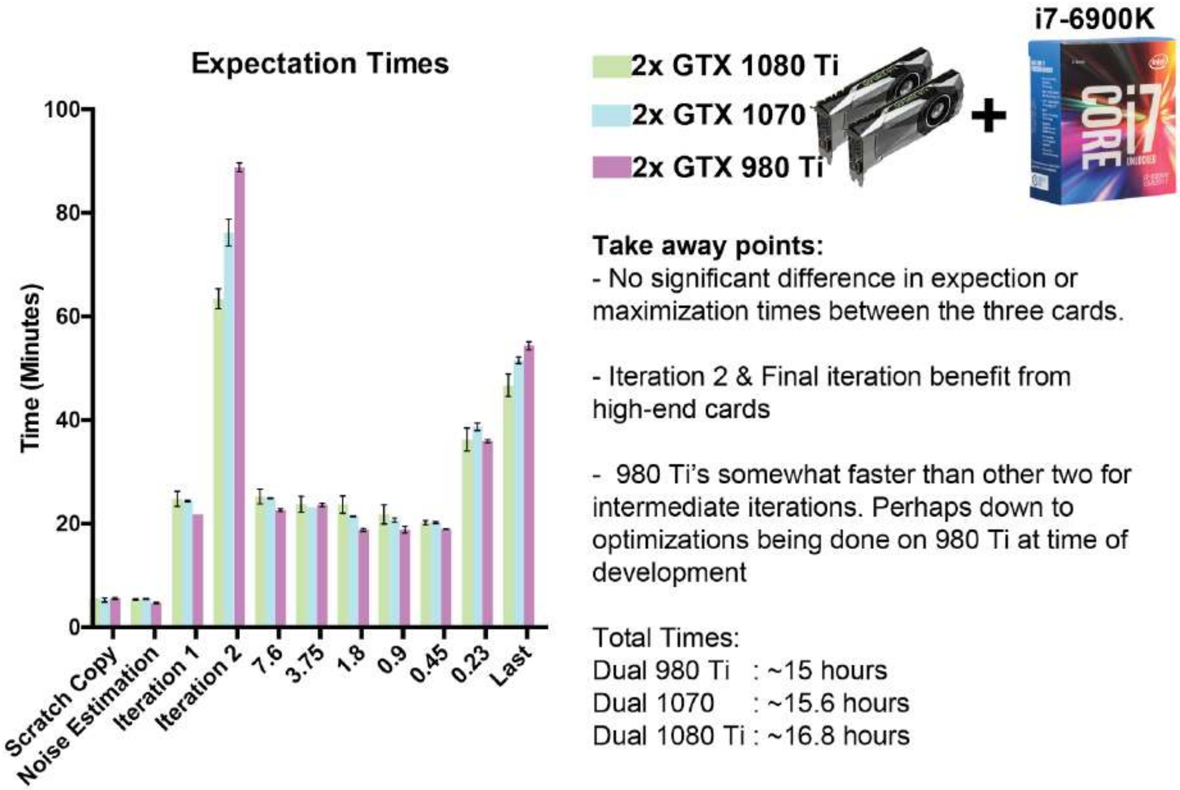
Two ways to alleviate this I/O (input/output) bottleneck within RELION exists, to either pre-read the particles into RAM or to copy the particles onto a faster SSD "scratch drive". We typically use the latter strategy since most desktop systems do not have system memory to house the particles as well as conduct the run, indeed when we tried to read the particles for the benchmark, we were greeted with an instant crash. SSDs tend to come in two flavours these days, the traditional SATA SSDs, and the newer NVMe SSD.



While the NVMe SSD offers a small improvement in iteration times compared to the SATA SSD, it is clear that the majority of the I/O bottleneck is alleviated when the particles are copied onto a SSD of any description. Nonetheless, the price of NVMe drives are not significantly higher than that of SATA SSDs, and since TR1950X has dedicated PCIe lanes for the newer standard, we would recommend getting NVMe drives.

CPU Bottlenecking

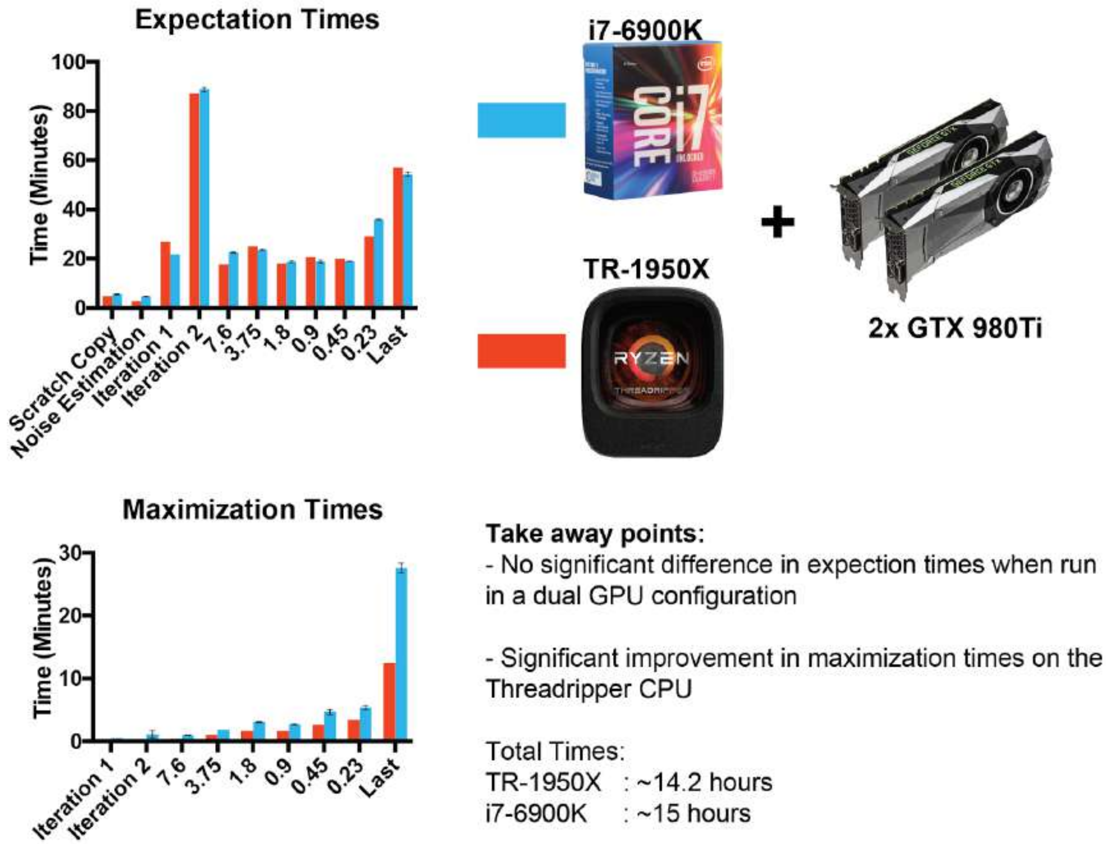
While in the community there is a strong impression that improving the GPU is the utmost important factor when building a workstation, Our recent benchmarks also reveal the importance of having a powerful CPU to accompany. We tested our “cost-efficient” workstation equipped with a 6900K CPU, by no means a weak CPU, equipped with three different types of GPUs: 980 Tis, 1070s and 1080 TIs in a dual configuration.



To our surprise we found that for the same run, there was little difference between the three graphics cards. The expected “on-paper” performance improvements were seen only on iteration 2 and the last iteration, but the intermediate iterations all saw similar timings, and if anything, in favour of the 980 Ti.

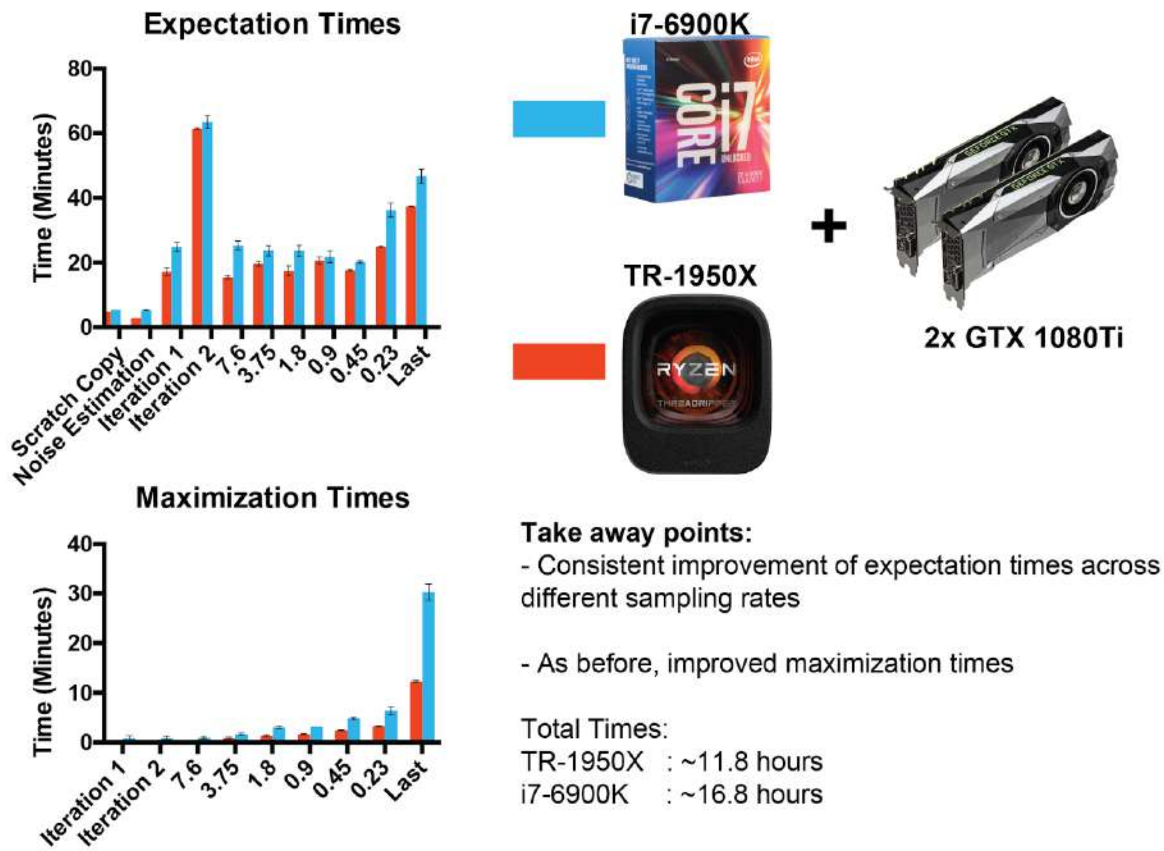
This shows that GPU accelerated RELION still has a large degree of CPU dependence, and with our i7-6900K, we are completely CPU bound and are unable to leverage the power of newer, more powerful GPUs for this benchmark.

Since there is an apparent CPU bottleneck for intermediate iterations of refinement, it would be natural to anticipate there to be an improvement in performance if we upgrade the CPU to something more recent, such as the TR-1950X.



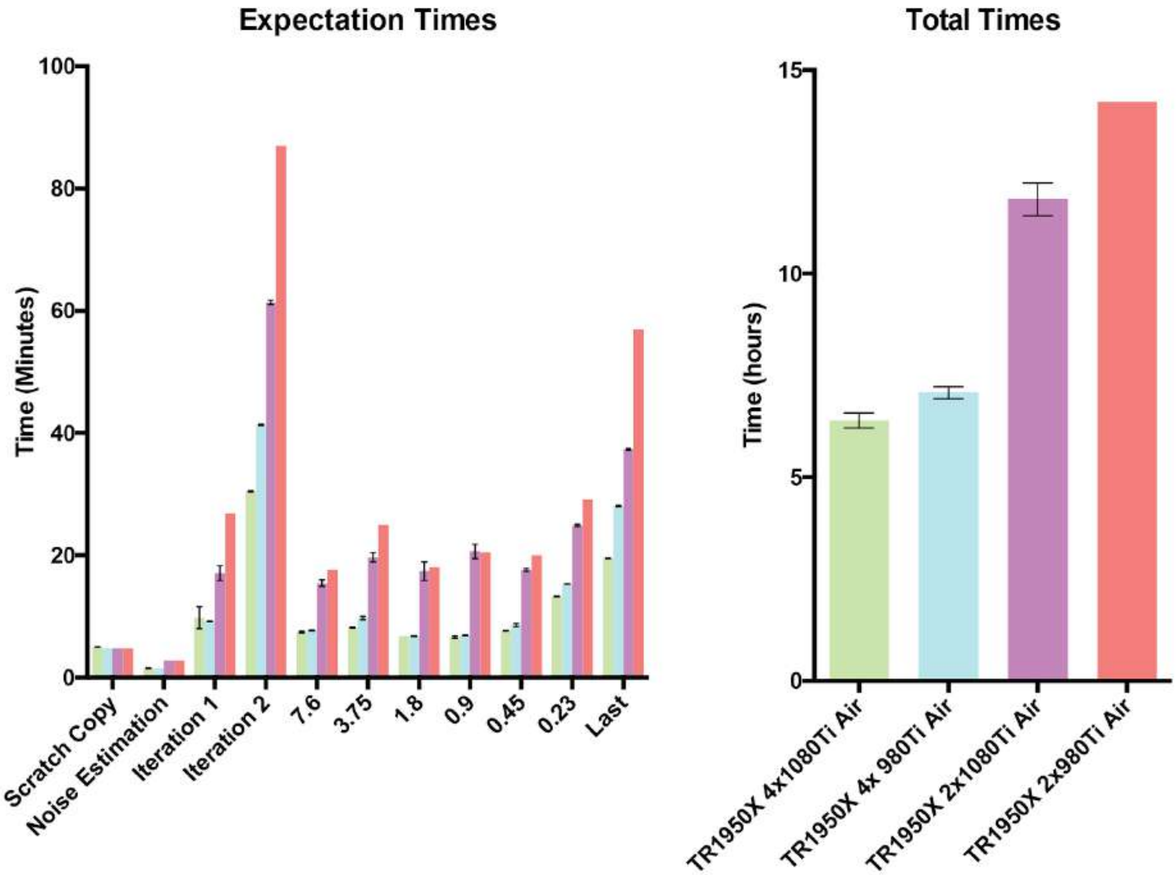
When we tested the TR1950X with two 980 Tis, we found that the timings for the expectation steps to be similar for both systems (expected since the GPUs are the same), however the TR1950X showed its power in the maximization steps, reducing the overall runtime by about 50 minutes.

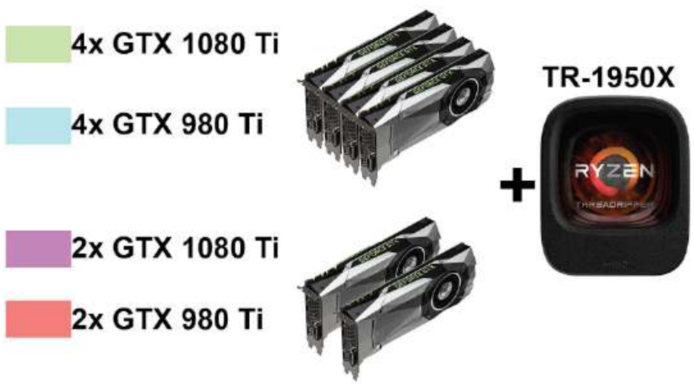
Now, it would be intriguing to see if the Threadripper can utilize more horsepower of the 1080 Ti.



It is clear that the Threadripper is somewhat better at leveraging the power of the new 1080 TI's, and compared to the 6900K system which did not benefit from the 1080 Ti, the Threadripper system sees improvement in the expectation steps when equipped with 1080 Tis. This brought the runtime on the TR1950X system with two 1080 Tis down to just under 12 hours, improving ~2 hours compared to having 980 Tis, and a 5 hour improvement compared to an equivalent 6900K system.

While it is clear that the Threadripper CPU is suitable for dual-GPU systems. In particular if you want to use the power of high-end graphics cards, it's real allure comes from its many PCIe lanes. We therefore tested quad-GPU configurations using either 980 Tis and 1080 Tis, and we were pleased to find that the performance scales well.





Take away points:
- Quad-GPU scales well on the Threadripper system

- As before biggest difference in intense iterations for GPU load

Total Times:
Quad-1080Ti : ~6.4 hours
Quad-980Ti : ~7.1 hours
Dual-1080Ti : ~11.8 hours
Dual-980Ti : ~14.2 hours

With a Quad-1080Ti system, we were able to run the benchmark in about 6.4 hours, and 7.1 hours for the Quad-980 Ti. Having four GPUs benefits the benchmark under all expectation steps at different samplings, however as before, the performance gains of the higher end cards are most noticeable for iteration 2 and the final iteration.

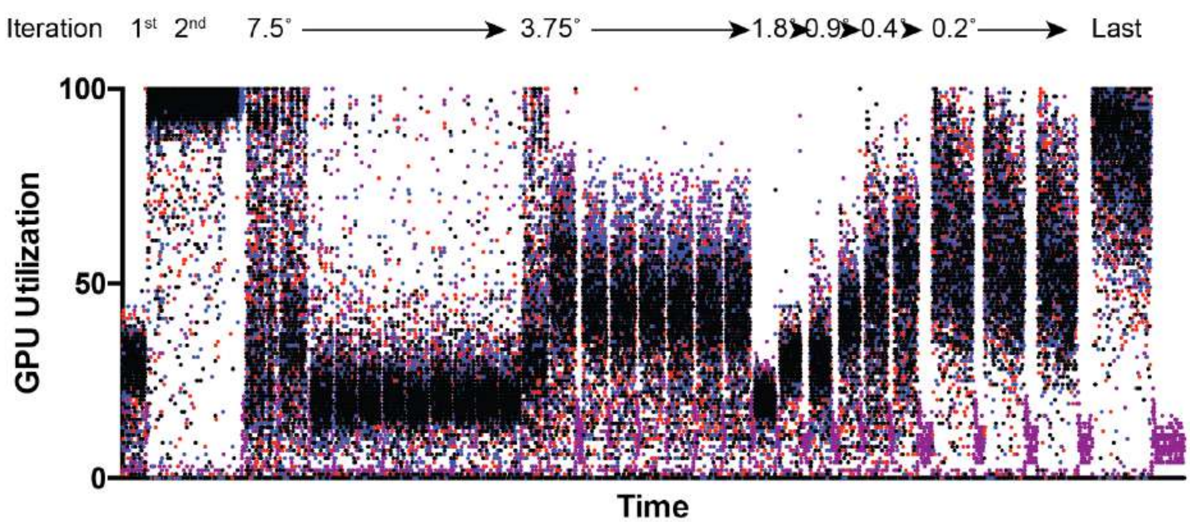
This shows that even the TR-1950X struggles to keep four graphics card entertained and occupied fully during the run.

Discussion

It is evident from our testing results that improvements in the RELION code has substantially improved both speed and scalability. In this report, we have shown that for ribosome refinement benchmark, depending on the setup, times can vary from upwards of 24 hours to 6.5 hours.

However, we have also highlighted that for a large proportion of the refinement benchmark, the GPU choice makes little difference on the iteration times. A deeper look into the GPU utilization reveals that the GPU utiliziation is only pinned at a 100% during the 2nd and last iterations - the iterations at which we observed benefits from using higher end cards.

In the other iterations, the GPU is not fully utilized, which means there is untapped resources on the cards, which is the reason why that iteration times do not change between 980 Tis, 1070s, and 1080 Tis.



We hope that future improvement in the implementation will be able to utilize this extra computing power in high-end graphics cards, but for now, if you are satisfied with the amount of VRAM on your recent generation GPU you will probably find the performance to be comparable to the highest end cards.

Benchmark methodology
We have run benchmarks using the cytoplasmic ribosome from *P. falciparum* deposited in EMPIAR EMD-10019. We have run numerous 3D auto-refinement runs since these are usually the most time-consuming step during processing. To minimize variability, all 3D auto-refinements were conducted with both --random-seed 0 and --perturb 0 for consistency. However even with these additional arguments, the total time for a 3D refinement varied substantially depending on the number of iterations that relion deemed necessary for the run (varying between 31 and 33 iterations in our hands). Therefore we have also presented the times in groups for a given angular sampling, and keeping the initial, 2nd, and last iterations in their own groups.

All tests were conducted using Relion 2.1.b1 and CUDA8.0, openmpi 2.0.2 or 3.0.0 (in our hands the version of openmpi had minimal impact on the iteration times). Temperatures and clock/memory frequency of the GPUs were monitored with nvidia-smi.

Appendix I: The price of silence

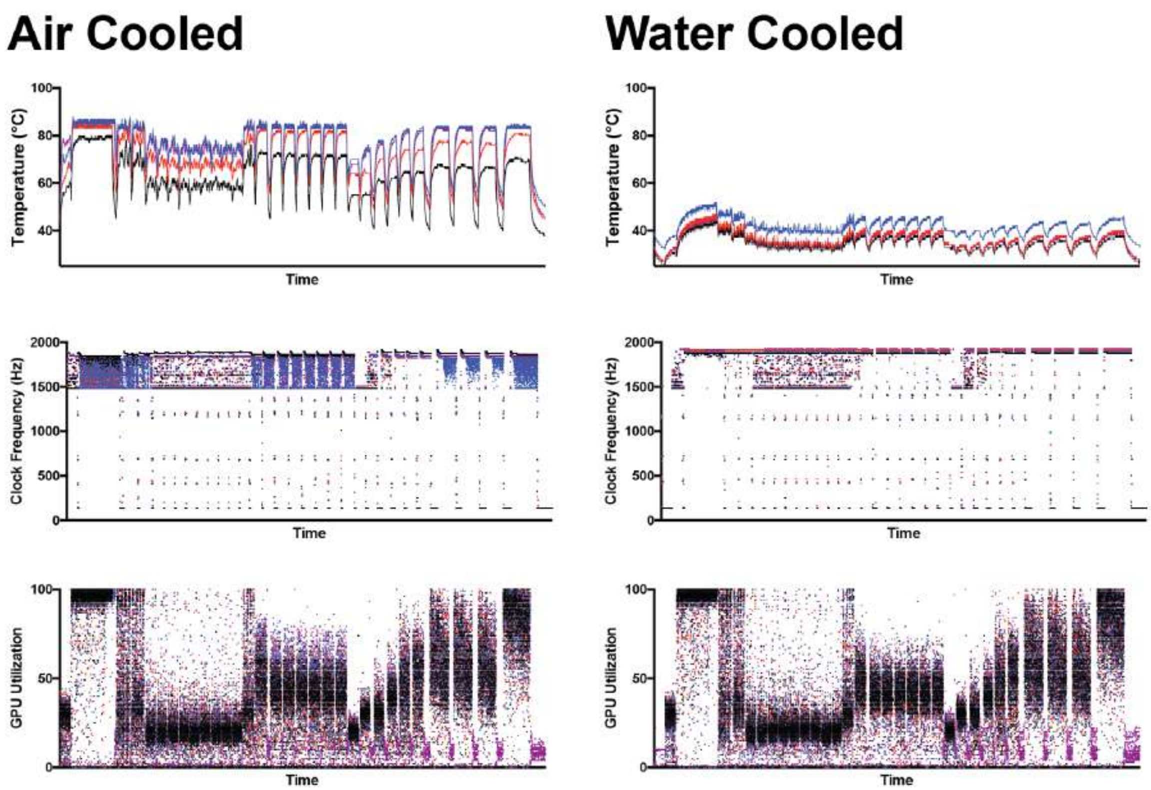
This section requires good understanding hardware to execute, we will not be responsible for damage caused to your hardware when attempting these modifications

One aspect that people have gotten to accept is the noise associated with these multi-graphics card workstations. As mentioned above, we recommend graphics cards with rear-exhaust "blower-style" coolers in order to minimize the potential for the whole chassis to overheat and potentially risk reducing the lifetime of the GPUs and other components within the system. The small, single, high-rpm fan on these cards can make a rather unpleasant racket and may attract the criticism of non-cryo-EM colleagues. However, if one is dedicated, there is a way to bring the noise level down.

One easy solution is to install the workstation in an isolated room with dedicated temperature control (for example, a server room) so that it is not a nuisance. However this is not an option for all, and having access to the workstation may be desirable since cryo-EM data is large and may involve physical drives being accessed by USB.

Another method is to cool the components with something more efficient - water. The principle idea behind water-cooling is simple. Heat from the heat generating components (CPU and GPU) are transmitted to water via nickel-plated copper block and dissipated to the atmosphere via a radiator. We take advantage of the large surface area of radiators to efficiently dissipate the heat into the atmosphere with relatively low-speed fans, thus minimizing the noise from the system. It is not unusual to reduce the temperature of a graphics card by 20~30 degrees using this method, allowing for further headroom for overclocking or increased lifespan of the cards.

In our system, we were successful in dropping the temperature of the GPU by almost 40 degrees. This allows the cards to maintain a higher clock frequency during the whole refinement, since on air the cards have to downclock frequently to keep their temperatures below the critical temperature. This in theory should translate to more consistent and faster performance, however, we must wait for further improvements in the code to actually reap the benefits of the higher clock frequency that can be maintained by cooling on water as described in the main text.



Counter-intuitively, a watercooled system actually houses more fans than an air-cooled system; but each fan spins at a lower rpm, lowering the overall system noise. Fan speeds are usually tied to the water temperature since there is no point in having the fans ramp up if the coolant temperature is low. To allow this, it is convenient to have a 2-pin water temperature sensor built into the loop which the UEFI BIOS can read-off and set up a curve based on these temperatures. We can say for sure that ASUS boards such as the Zenith Extreme has 2-pin headers and also allows for custom fan curves based on water temperatures. MSI boards also have 2-pin headers, but we have not tested whether the BIOS allows for a custom fan curve.


In our system, we cooled the TR1950X CPU and four GTX 1080 Tis with water. We have explicitly chosen to use 1080Tis with reference-design PCBs since when convereted to a watercooler, the reference-design 1080 Ti occupy exactly 1 PCIe slot. This allows for us to utilize the 10-gigabit ethernet card on the Asus board that was previously not installable due to the dual-slot cards. This is a feature of reference-design PCBs since they do not have a DVI port, which typically occupies the 2nd slot.

If you are not particularly attached to using ASUS boards or a 10 gigabit ethernet connectivity, there are 1080 Tis with waterblocks pre-installed such as the MSI Seahawk-X EKWB, or the Gigabyte Aorus Waterforce Extreme. Allowing one to preserve the warranty of the graphics card and save the hassle of installing a waterblock. One could also consider using the Asrock board with native 10-gigabit connectivity and come up with a custom fan curve (or leave them at a constant low-rpm).

To ensure we have enough radiator surface area to allow for near-silent operations, we have dedicated 240 mm of radiator area per heat-generating component with 2x 480 mm radiator and one 240 mm radiator. A quad GPU + CPU + 1200 mm of radiator space is definitely the upper end that most PC builders will encounter in terms of total restriction of flow; to ensure that we can get enough water flowing across all components, we have two D5 pumps in serial so that we have good flow and head pressure.

In our build we have used hardline tubing, mainly for aesthetics although it does provide a little bit more structural rigidity. It would be perfectly reasonable to use soft tubing which is much easier to handle. It is however that the correct type of fitting is used for the tubing type and size.

A drainage system is essential for a hardline system and definitely convenient to have for even a soft tubing system. We have plumbed a drainage system at the lowest point of the system, using a T-splitter, 2 dual-rotary male-to-male connectors, one ball-valve, a stop plug, and some soft tubing.

The Corsair 900D Supertower is an excellent case to house such a system. This case can house a 480 mm radiator (60mm thick) in the top and bottom, with further space for a 240 mm radiator (40 mm thick) in the bottom compartment.

Full List of Watercooling Parts:

Waterblocks and Radiators

1x EK-Supremacy EVO Threadripper Edition - Nickel
4x EK-FC1080 GTX Ti - Nickel
4x EK-FC1080 GTX Ti Backplate - Black
1x EK-FC Terminal QUAD Semi-Parallel
2x EK-CoolStream XE 480 (Quad)
1x EK-CoolStream PE 240 (Dual)

Pump, Reservoir and Coolant

1x EK-XTOP Revo Dual D5 PWM Serial - (incl. 2x pump)
1x EK-RES X3 250
3x EK-CryoFuel Clear Concentrate 100 mL
1x EK-ATX Bridging Plug (24 pin)

Angled Fittings, temperature sensor, and drainage system

6x EK-AF Angled 90° G1/4 Black Nickel

6x EK-AF Angled 2×45° G1/4 Black Nickel
1x Temperature sensor inline G1/4
2x EK-AF Extender Rotary M-M G1/4 - Black Nickel
1x EK-AF T-Splitter 3F G1/4 - Black Nickel
1x EK-AF Ball Valve (10mm) G1/4 - Black Nickel
1x EK-ACF Fitting 10/16mm - Black Nickel
1x EK-Tube ZMT Matte Black 15,9/9,5mm

Hard tubing, fittings, and bending tools
14x EK-HDC Fitting 16mm G1/4 - Black Nickel
4x EK-HD PETG Tube 12/16mm 500mm (2pcs)
1x EK-HD Tube D.I.Y. Kit 10&12mm
1x Heat gun

Soft tubing and tools (alternative to hard tubing)
14x EK-ACF Fitting 10/16mm - Black Nickel
2x EK-DuraClear 9,5/15,9mm 3M RETAIL

Fans and fan splitters
2x 5-way PWM fan splitter (AKASA)
5x ML120 120mm PWM Premium Magnetic Levitation Fan — Twin Pack

Links to items in order appearance of above
https://www.ekwb.com/shop/ek-supremacy-evo-threadripper-edition-nickel
https://www.ekwb.com/shop/ek-fc-1080-gtx-ti-nickel
https://www.ekwb.com/shop/ek-fc-1080-gtx-ti-backplate-black
https://www.ekwb.com/shop/ek-fc-terminal-quad-semi-parallel
https://www.ekwb.com/shop/ek-coolstream-xe-480-quad
https://www.ekwb.com/shop/ek-coolstream-pe-240-dual

https://www.ekwb.com/shop/ek-res-x3-250
https://www.ekwb.com/shop/ek-xtop-revo-dual-d5-pwm-serial-incl-pump
https://www.ekwb.com/shop/ek-cryofuel-clear-concentrate-100-ml
https://www.ekwb.com/shop/ek-atx-bridging-plug-24-pin

https://www.ekwb.com/shop/ek-af-angled-90-g1-4-black-nickel
https://www.ekwb.com/shop/ek-af-angled-2-45-g1-4-black-nickel
http://shop.aquacomputer.de/product_info.php?products_id=2291&language=en
https://www.ekwb.com/shop/ek-af-extender-rotary-m-m-g1-4-nickel-1
https://www.ekwb.com/shop/ek-af-t-splitter-3f-g1-4-black-nickel
https://www.ekwb.com/shop/ek-af-ball-valve-10mm-g1-4-black-nickel
https://www.ekwb.com/shop/ek-acf-fitting-10-16mm-black-nickel
https://www.ekwb.com/shop/ek-tube-zmt-matte-black-15-9-9-5mm

https://www.ekwb.com/shop/ek-hdc-fitting-16mm-g1-4-black-nickel
https://www.ekwb.com/shop/ek-hd-petg-tube-12-16mm-500mm-2pcs
https://www.ekwb.com/shop/ek-hd-tube-d-i-y-kit-10and12mm

https://www.ekwb.com/shop/ek-acf-fitting-10-16mm-black-nickel
https://www.ekwb.com/shop/ek-duraclear-9-5-15-9mm

http://www.akasa.com.tw/search.php?seed=AK-CBFA03-45
http://www.corsair.com/en-eu/ml120-120mm-premium-magnetic-levitation-fan-twin-pack

Appendix II: PCIe 16x vs 8x

Devices such as graphics cards use PCIe lanes to communicate with the CPU. The number of PCIe lanes will differ for each CPU and how those lanes are distributed across the board depends on the motherboard manufacturer. Relevant for RELION, NVMe SSDs and GPUs depend on PCIe lanes for communication to the CPU.

You will likely find typical system configurations to have 16x or 8x PCIe lanes going from the CPU device to the graphics card. Since typical CPUs do not have enough lanes to allocate 16 lanes for all four graphics cards, you will typically find configurations like 16x/8x/16x/8x.

In our tests, we have compared the performance of a 16 lane connection vs an 8 lane connection using our cost efficient workstation by bringing the generation speed down in order to simulate a 8 lane connection (16 slot PCIe gen 2 = 8 slot PCIe gen 3). We have found that there was no significant drop in performance when the bandwidth of the graphics cards were halved and so would be comfortable to recommend systems with a 16x/8x/16x/8x connection. We envisage 16x/8x/8x/8x to perform reasonably well, as long as lanes are not going through a switch as previously described.

Appendix III: Overclocking

Our early tests indicate that in particular on the Threadripper platform, overclocked RAM has some impact on the runtimes of RELION (in our case from 2133 MHz → 2400 MHz) however stably overclocking these kits while maintaining a low CAS latency is proving difficult and so we have been unable to test further overclocks.

GPU overclocks give a small boost in performance, however given the low utilization of the GPU during refinements, the gains are not as great as one would expect.

Appendix IV: General Building Notes

We do recommend testing the system outside of the case prior to assembly to identify any faulty components since re-plugging components is significantly easier prior to mounting.

Threadripper CPUs are bigger than any other CPUs I have dealt with in the past. We have found that a fairly generous serving of thermal paste can help the thermals of this CPU, and so we would recommend applying more than you would traditionally.