

Supply chain strategy for research computing throughout the pandemic.

Jake Carroll

The University of Queensland Research Computing Centre, Australia.

cto@rcc.uq.edu.au



Acknowledgement of **Country**

The University of Queensland (UQ) acknowledges the Traditional Owners and their custodianship of the lands on which we meet.

We pay our respects to their Ancestors and their descendants, who continue cultural and spiritual connections to Country.

We recognise their valuable contributions to Australian and global society.



Preamble.

- I wrote this talk because the pandemic changed my world view.
- I had to think carefully about how to change our practices in order to help researchers continue their eResearch intensive work.
- I care a great deal about providing world class infrastructure for scientific research, but it is irrelevant if we cannot obtain that infrastructure.

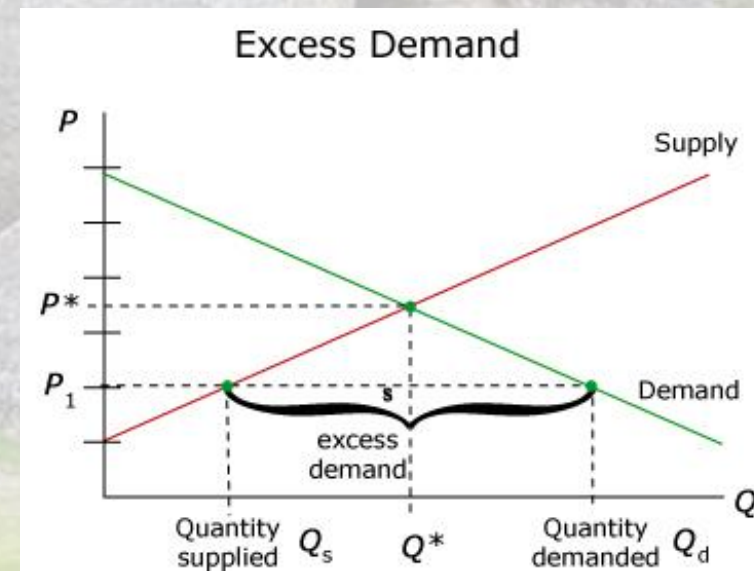
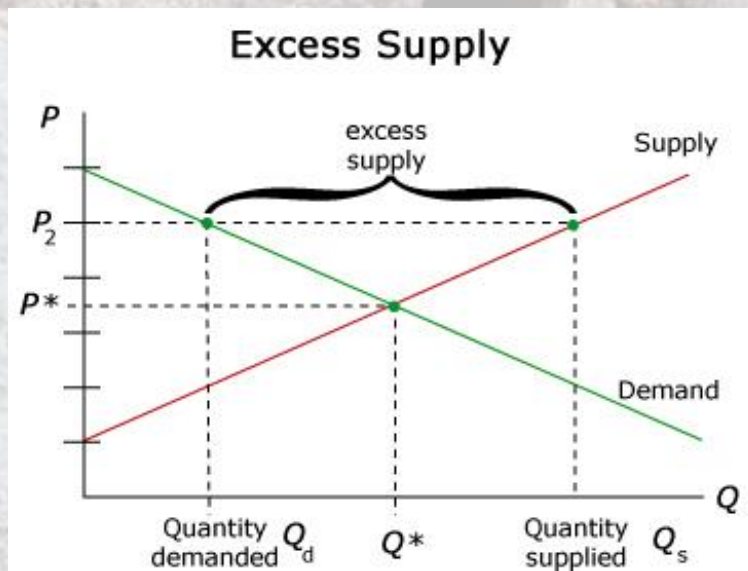
Today.

- This talk is about how the UQ RCC managed the impact of the pandemic upon our ability to provide scientific computing infrastructure.
- If you take nothing else away from today:
 - Cooperation, compromise, consideration, pragmatism and patience are more important than one might think.
 - We need to get out of our own heads, processes, institutions.
 - We need think globally in order to succeed.

C O N T E X T

What chip shortage?

- Economics 101: In simplest terms, the current chip shortage is due to strong demand and no supply.
- Goes back to COVID-19 lockdowns in the second quarter of 2020.
- Demand for work-from-home technology increased exponentially and car makers found themselves competing for the semiconductor capacity located in Asian foundries.



- Adding to the problem, downstream operations in South Asia were adversely impacted by the COVID-19 Delta variant, creating further bottlenecks in the supply chain.
- Malaysia performs many “back-end” operations such as chip packaging and testing.
 - More labour intensive than wafer fabrication processes, so activity was more easily affected by public health measures.
- Then it got worse: drought in Taiwan struck. Fab is thirsty business!

TSMC trucking in waste-water for purification and re-use to tackle their water shortage issues. → Maybe they always should have been, drought or otherwise?



The cancellations begin...

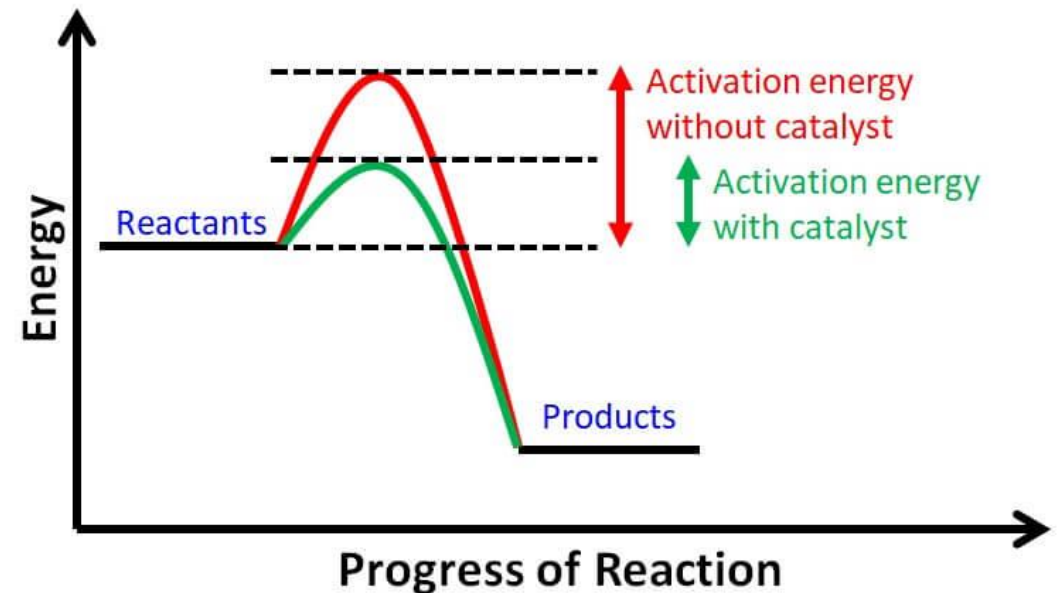
- At the beginning of the pandemic, car companies cancelled orders, but as production ramped up again towards the end of 2020, there was no semiconductor supply available.
- This was compounded by demand increases, particularly at the higher end of the car market, **as low interest rates were aiding affordability.**



ORDER CANCELLATION REQUEST

Reactants and products.

- While the COVID-19 pandemic was the **initial** catalyst for the chip shortage, **systemic structural factors** are also part of the picture.
- Many years of status quo lead up to a failure.
- The car industry is changing, with a major shift towards automation and electric vehicles. These require yet more chips, causing further strain on an already stretched industry.
- We were left to pick up the pieces and try as best we could to buffer and absorb the shock of this situation.



Global tensions and events.

- The war in Ukraine had global reach in supply chain disruption, focus, logistics and determinism of supply routes.
 - Energy costs skyrocketed in the EU – semi-fab + dev costs went up.
- Inflation in the United States meant semi-conductor development cost knock on effects.
 - Supply-cost effect, COGS effect.
 - Massive cost of labour impact, stateside (*contrast Australia – with wage stagnation...*)
- The geo-political tension between China and Taiwan complicated matters even further.
- We now know this has gone as far as international embargos, “do not sell” administrative rulings and silicon “will not supply” circumstances with certain vendors.



Trade Wars

Raw Material Shortages

Covid-19 Pandemic

Drought in Taiwan

Global Chip Shortage

PCBs/motherboard shortages for our servers hit here

March
Renesas factory fire



July
Nittobo factory fire



Three major vendors could no longer guarantee we could get PSUs for our servers on time, after this.

February

- Earthquake causes blackout at Renesas factory
- Texas winter storm suspends Samsung, NXP and Infineon plant operations



October
AKM factory fire



Memory (RAM) shortages for our DDR4 + DDR5 server platforms in HPC

2018

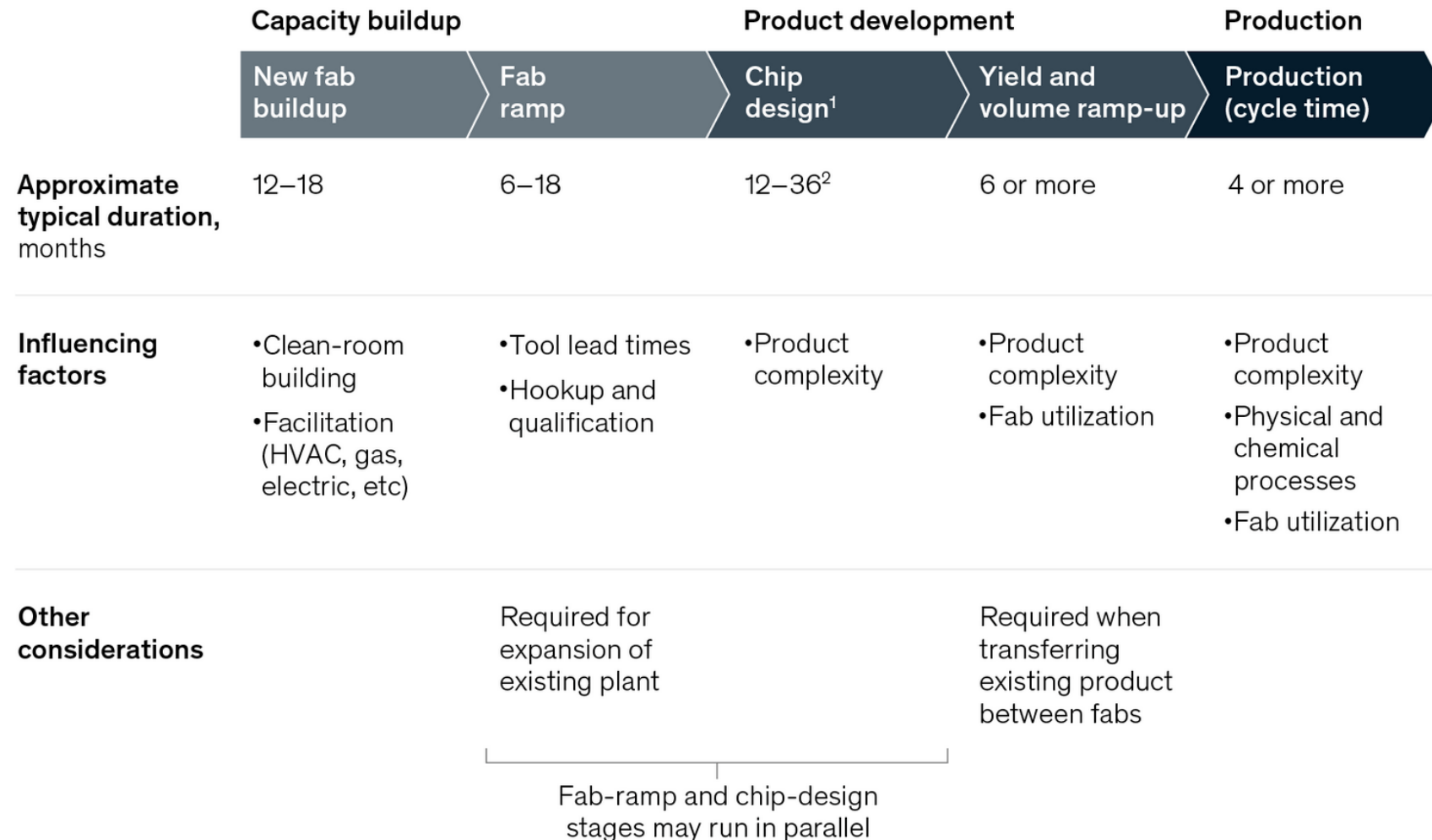
2019

2020

2021

We cannot simply “turn on a new fab” overnight, or switch to a new manufacturer. It can take a year or more to turn the ship around.

Semiconductor development and production timelines





Four thought processes we used to work through this.



1: A global view.

The **broadness** of our **perspective** needed to change.

Instead of worrying about the RFP/RFQ/Tender, we needed to shift our focus to global events to understand how the market might respond – such that we would not ask anything impossible of it, wasting the time of our suppliers and our own.



2: Compromise.

“But I want <insert component here> delivered on time, this quarter, with these attributes, Jake!”

Me: “That’s nice. How about no? You need to think a little more carefully about how the world has changed. If we ask a supplier to deliver things they cannot – what makes you think they’d be fine with that? We need to ASK what is possible and work towards the middle. Not DEMAND the earth.”

“Drop dead dates, contractually binding delivery schedules! That’ll help us get this stuff and some certainty!”

Me: “Actually, it’ll get you the exact opposite. The harder you push, the less likely and less obliging your hardware industry partner will be if you try to force them down a path of contractual impossibility given the uncertainty, they themselves face. Remember what I said about thinking more carefully?”

“We want EXACTLY that specification!”

Me: “Cool – know that you may never get anything at all then. How about we work on some compromise?”



3: Knowing our place.

Our place, in space.

- Unfortunately for us, we share some exclusive company.
- We are at the bottom of a very large market. We are in the same realm as hyper-scalers.
- UQ RCC buys thousands of CPUs and HDD/SSDs/tapes at a time. They buy **hundreds of thousands** at a time, if not **millions**.
- We are at the mercy of the supply “tide” – we take what is left from our bigger brothers and sisters after they’ve had their say and have eaten.
- Unfortunately, the scale we now operate at means we are lumped in the same supply bin as they are.
- We must plan accordingly.



UQ RCC now deals in 10PB “chunks” multiple times per year.

Our place, in space – second order effects.

- It has a second order effect.
 - If the hyperscalers decide on the use of a technology, because we live in the same ecosystem they do – we get pulled along in the rip-tide with them.
 - Example:
 - Let's say a hyperscaler decides a specific type of tape isn't competitive to them – they will not use it. They will signal to the industry, consortia and the technology committees that they were not happy with the economics or price/performance ratio.
 - We watch the tide rise – we see the product they disapproved of take the hit – and we watch what everyone else thought was the “right and best commodity” suddenly become unpopular, expensive and unsupported.
- Like it or not, at that point, we follow them. They may have decided on a different technology, which, because of scale, then perverts the market in choice, scope and cost.



4: Accepting that we will
need to micro-manage.

Everybody's fight, together.

- Everyone must get their hands dirty.
- The CTO that fails to wade right in and get completely involved because she/he/they are “*above that micro management*” is the CTO that fails to get the outcome.
- Ordinary times: “*The procurement office will take care of this for you*”
- Now: “*I will wait up until 11PM for the delivery notification so we know it got into the country so we can have it at the data centre on time. Major issues in customs. Needs hands on. I will be on the conf-call at 1AM with you in New York to talk options and contingencies*”

Playing out every scenario. Looking under every rock.

- We spend months working through the possibilities of the market, events and it's impact on what we might do. This is deeper than procurement. This is market intelligence and global economic conditions understanding in order to even have a shot at making it all work for research. **The engine room.**



Multi-region sourcing.

- Worked for us on three occasions but required us and our hardware partners to be incredibly well aligned, diligent, careful and in constant communication around the clock.



Organisational, Behavioural and socio-cultural factors

- **Stop siloing your teams.**
 - Smart governance tzar's will see the procurement, legal and finance people as one organism to do the collective WorkToBeDone™ and achieve the outcome
 - ...even if those teams see themselves as very separate functions, both organisationally, culturally and behaviourally.
 - Help these teams break down barriers, work together to get the outcome you need – not as separate business units...



Legal, contract, law.

- Spend time with your legal team.
- Help them understand the nuance of the situation.
- Read them into nature of the silicon shortage crisis.
- It will help you enormously, when context and sensitivity is required.
- They will “get” your struggle and the sensitivities of the other party, too.
- Nobody gets backed into the corner.
- Parties have the (reasonable) option to cool it off and walk away with little loss apart from the administrative overhead to have gotten that far.
- Start being more reasonable. Stop being so defensive. The working mechanics of “risk” needed to change as the ground shifted. Everything is a risk. Everything is unsure. Work with the inputs you have and behave accordingly.

Partners. Less us and them.

- Working closely with your industry partners and providers.
- They **don't** want to lose either.
- They **don't** want to disappoint you. Reputational damage is too great.
- The game of “us and them” is long dead.
- Stop treating your vendors adversarially. Yes, they are there to make money – but they are also part of the team that helps you deliver capability.
- Start to understand the ecosystem that feeds **all** of it.
- Ask deeper questions about their pressures on their own supply chains.
- Understand the full spectrum so you can make more educated decisions that take into account their perspective too – lest you not get much at all.
- Ditch the hubris. We **don't** always know best. They might know more.
- Play “*how might we be wrong?*” more often...

The amazing people I get to work with.



Ms Sarah Walters



Dr Marlies Hankel



Dr David Green



Mr Ashley Wright



Mr Irek Porebski



Mr Jake Carroll

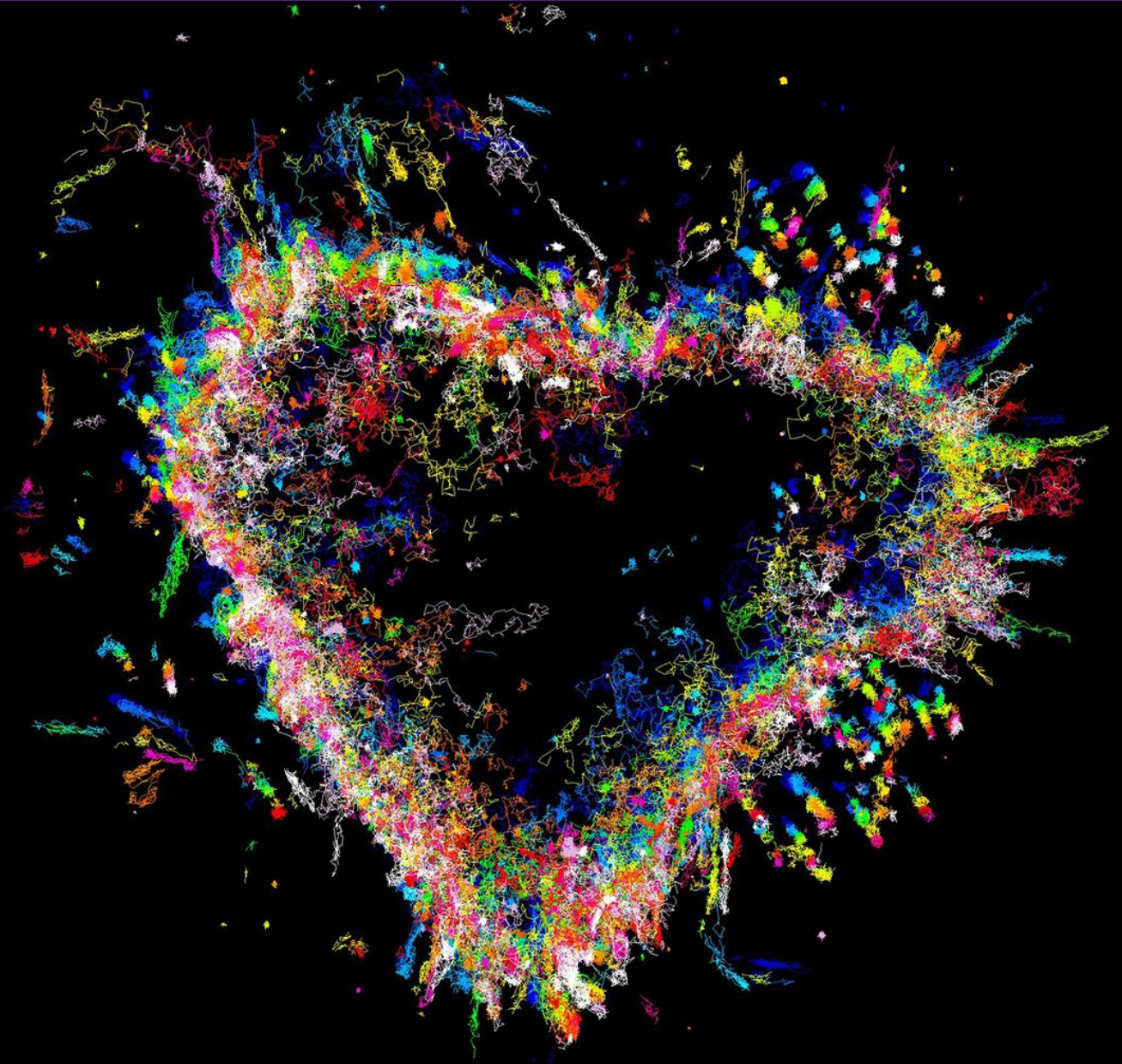


Mr Owen Powell



Professor David Abramson

Thank you.



Question time.



Image and Credits

Slide 2: *Hippocampal neuron labeled for actin (orange) imaged by STORM and synapsin (blue) imaged by TIRF.* Credit: Christophe Leterrier, NeuroCyto Lab, INP CNRS-Aix Marseille University, Marseille, France. Retrieved from <https://www.technologynetworks.com/neuroscience/articles/what-is-super-resolution-microscopy-sted-sim-and-storm-explained-328572> on 26/06/2022

Slide 3: *A STED (stimulated emission/depletion) micrograph image revealing actin (magenta) and microtubules (cyan) of a young dissociated hippocampal neuron.* Credit: K. Jansen and E. Katrukha, Kapitein Lab, Molecular and Cellular Biophysics, Utrecht University, The Netherlands. Retrieved from <https://www.technologynetworks.com/neuroscience/articles/what-is-super-resolution-microscopy-sted-sim-and-storm-explained-328572> on 26/06/2022

Slide 16: *iSIM (instant structured illuminated microscopy) image of primary rat cortical neurons aims to understand the nanoscopic organization of synapses using stains for: MAP2 (blue) as a morphological marker; the presynaptic protein bassoon (red) as an active zone marker; and SV2a (Synaptic vesicle glycoprotein 2A, green) as a marker of synaptic vesicles.* Credit: Deepak P. Srivastava, Director of the Wohl Cellular Imaging Centre at King's College London, UK. Retrieved from <https://www.technologynetworks.com/neuroscience/articles/what-is-super-resolution-microscopy-sted-sim-and-storm-explained-328572> on 26/06/2022

Slide 28: *Kasula, R., 2017. PhD Student, Image Competition, QBI, The University of Queensland 2017.*