In the field of astrophysics there are many outstanding issues that can be with the aid of computers. As space exploration is currently conducted by observation and prediction on how bodies will behave following certain fundamental laws, computers aid in this research by returning solutions more efficiently and with greater accuracy. Discovering habitable planets or understanding the future of our sector in space are key areas of interest which are assisted by computing technology by measuring minute dips in a star’s brightness and by observing and predicting how galaxies evolve and change over time. The movement of galaxies and other heavenly bodies through space is a subsection of hydrodynamics which relies on continuous equations.

An equation for describing movement of this nature contains many variables including change in momentum, pressure, volume and more. A good example of how computers can help researches in by solving these equations which can also be using parallel computing techniques. Algorithms are constructed which involve differing values for variables which can be simulated concurrently. An algorithm with one set of variables can also be run over several processors simultaneously to improve the speed of delivery even further. Multithreading can lead to situations which create more problems when implemented inefficiently such as livelock where a programs threads are communicating so often that useful behaviour ceases to function or deadlock which is a condition where each thread is waiting for the other to release a shared resource so it may finish its task.

Dr Acreman raised many interesting points about the use of and importance of computers in his field of study. An explanation of how continuous equation can be computed revolved around the use of grids for calculating star formation and like phenomena. A section of space is represented by a grid or an array and the variables to be used are stored within sectors of the entire grid. The changes are between sections over time can be calculated to provide an estimation of the real world conditions. The closer to real world values are. Given more accurate data, the results of the simulation will be more accurate however Dr Acerman demonstrated just how vast the storage needed for relatively simple situations can be. 40 bytes of data can be used to represent simple forces operating on a single sector of a grid, given a grid with one thousand sections in each three-dimensional direction approximately 40 GB of data make up one instance of a continuous equation converted into a grid. Some aspects of these calculations will scale exponentially such as the interaction of gravity between gasses as well as the need for the equations to be run several times, with differing variables or time-slice intervals. Dr Acerman’s research grants him access to some impressive computer equipment, with a brief outline of of the specifications given with consideration given to storing data with cloud services.

One algorithm that successfully reduces the memory imprint needed for calculations of this type as well as giving accurate results is to use an adaptive mesh interface. An outline of the algorithm as described in the paper local adaptive mesh refinement by Marsha Berger and Joseph Oliger.

1. The computational domain is covered by a Cartesian grid

2. determine criteria based on user input

3. label or tag grid cells based on supplied or extracted criteria

4. apply a more accurate grid on top of the tagged sections of the grid

5. for a given level of refinement, pass the variables contained within this section to an integrator

6. the integrator advances the sections in time

7. a correction procedure is applied to transfer information from more refined sections of the grid to less refined (or coarse) sections. This insures quantities are preserved when transferring between high and low details sections.

8.if a section of the grid no longer needs to be monitored at a high resolution, drop back to a broader layout for that section or enhance further if required.

This algorithm has been used to reduce a uniform grid containing 10^15 sections down to less than 150,000 cells to model a giant molecular cloud

Richard l. Klein, (1999) Star formation with 3-D adaptive mesh refinement: the collapse and fragmentation of molecular clouds

Mr Charles Ewen from the Met office gave a presentation on of the work conducted by the Met Office. The Met office is one of the largest data houses monitoring and predicting weather worldwide so it was very interesting hearing about the working environment, challenges overcome and huge collaborative effort between the scientific and business sectors worldwide. Out of the information presented to us about the computing capabilities controlled by the Met Office, one of the facts that I find most interesting to comprehend was that their supercomputer uses more power than the whole of the City it is based in. This helps to visualise the sheer scale of the types of calculations and simulations needed to provide Scientists with information about climate change as well as weather forecasts for the public and private organisations.

Despite this impressive computing power, the importance of even more powerful hardware was conveyed to us throughout the talk and in follow up lectures. Exscale supercomputers will, when developed, will provide even more outstanding performance in terms of what can be achieved. As the computing power progresses to keep up with the sheer flow of information harnessed so to do the communications networks. 5G networking will assist the Met Office with communicating with many times more devices and systems as the internet of things begins to become more commonplace. This could even allow for the monitoring of ‘micro-climates’ according to our guest speaker.

The met office performs Numerical Weather Prediction for its clients and models. This process can be defined as

1. obtain equations to describe the flow of fluids

2. programming numerical methods from equations

3. run with instances of different parameters

4. combine results and apply on a specific domain

5. integrate the result from the initial conditions

the initial equations involve, for example, the Conservation of momentum for 3d wind acceleration, ideal gas law and the conservation off mass. Depending on the needs of the computation and the location of the prediction, different emphasis can be placed on the initial equations as an airport may be more concerned with wind speed, while hurricane detection may be more concerned with air temperature. The equations used for Numerical Weather Prediction produce accurate results however it is debated in some scientific circles on the nuances of what is and isn’t conserved. From a computing perspective, a relatively simple equation for determining wind acceleration in one dimension affected by pressure involves calculus which is difficult to compute due to the the nature of how computers perform mathematics. This is why numerical methods are needed. By converting spatial and temporal derivatives into algebraic equations such as finite difference based on Taylor series model output at small scales can be more accurate. The advantage of parameterizing the prediction is that effects which take place on a scale which is currently to small to represent accurately can be represented to a respectable standard. Some of these effects would be better monitored with the proposed benefits of exscale computing such as the effect of radiation, pollution or even the Earths surface. This part of the algorithm shows the most room for improvement. The domains may be modelled in 2 or 3 dimensions such as river surface as air pressure or the progression of a thunderstorm. These may be represented in hexagonal, triangular or spherical grids. Subsections of grids can be analysed for greater accuracy or different terrain. Finally, the initial conditions and boundary conditions can be contrasted.

There are other algorithms used for prediction of the weather, the advantage of Numerical Weather Prediction in this case is that it takes advantage of the vast amounts of data the Met Office receives from around the world. One way to improve this algorithm would be to take previous predictions and observed weather and feed this data back into future predictions in a way which could adjust for discrepancies found to improve future predictions. As the Met Office and society in general places itself in a position to benefit from increasing amounts of data sending and sorting this data will become of paramount importance

Jason Knievel, Numerical Weather Prediction (NWP) and the WRF Model (2006)

A.A. Fomenko, Numerical Methods for Weather Forecasting Problems (2000)

Digital information is an increasingly important resource in the modern world. Companies depend on keeping accurate records as part of their obligation as a business and many modern companies are primarily built on the secure storage of this information. There is also a growing market for sharing resources in a similar manner. As individuals and companies of all sizes seek to make more use of the computing power at their disposal, datacentres are increasingly becoming important. Aside from the reduced cost of a project to have some elements conducted or stored on another site, data warehouses that conform to data protection laws can give an organisation one less process to monitor when all is conducted correctly. Sharing information across the globe is another benefit of entrusting data to a centralised computer system which is essential for science business and even entertainment. Storage as a service offer these and many other benefits that solve the issue of data explosion and security in modern society.

Care must be taken to mitigate the drawbacks of centralising data such as performance, traffic and data latency. As modern networks are configured to behave as a single device scaling and storage becomes an issue. data needs to be stored about the data held which may reduce the overall available space for data as well as resources as some proportion must be spent on monitoring existing data. One way this is addressed is by using different storage mediums witch balance attributes such as cost and speed, combined with algorithms which transfer less accessed data to slower sections of storage which increases overall efficiency of the data centre.

There are many different designs and architectures for data warehouses as well as the Storage Area Networks they make use of. To speed up the read and write speed as well as reduce errors when backing up or transferring data, RAID systems are implemented. RAID is a system of redundant independent disks and there are several different ways in which these are written to to solve the similar yet differing problems which arise when writing digital information. These roughly follow the same outline which can be described as

1. define system requirements and select appropriate RAID rules

2. calculate space on drives needed for data and parity information

3. write files to drives

4. if there is redundancy, restore data using raid implementation

A raid implementation may strip or split data across multiple drives. A simple implementation of this would be RAID 0 which is easily configured however offers no real protection against data loss although performance is increased and data from different sources can be combined easily. Another simple implementation is RAID 1 which is probably what most people think of in terms of backup, copying the data exactly as is to another drive. This insures that information read from the disk array is reliable however the speed of faster drives will be held back by the slower ones. In the talk given by David Barker, RAID 5 was highlighted as efficient. RAID 5 uses bitwise logic and stores data about how the information is divided amongst the discs as parity to distribute parts of the same data across multiple drives. The bitwise logic can determine which information is missing if a drive fails and this also increases write speed as it happens across multiple part of the disk array. In actual use, many raid systems are used alongside each other and the selection of which raid architecture is to be used can be carefully calculated based on metrics such as space efficiency, failure rate read or write performance as well as fault tolerance outlined here.

Cloud storage is reaching a more mature stage of development and as with most technological improvements will likely run alongside more traditional data warehouses which have been around for a while. Great improvements have been and will continue to take place using software defined methods, however new storage mediums are currently being developed such as glass and DNA which shows great promise in handling data for future needs.

Chen, Peter; Lee, Edward; Gibson, Garth; Katz, Randy; Patterson, David (1994). "RAID: High-Performance, Reliable Secondary Storage"

Dawkins, Bill and Jones, Arnold, (2006) "Common RAID Disk Data Format Specification

Computer systems can be found at the very heart of modern infrastructure. Power stations, air traffic and finance are some of the key sectors which rely on computers and thus are often targeted by hackers and other ill-meaning organisations. Cyber security is becoming less of a personal or business concern as entire nations depend on the smooth operation of their computing resources, as well as the need to shield sensitive information from people who should not have access. The threats from cybersecurity are increasingly becoming more varied and often more sophisticated although the reality of large hacks being perpetrated by small groups or even individuals remains. Although most computer systems contain passwords that can not be cracked in a realistic timeframe by even the most powerful computers around today, human engineering remains a viable option for hackers to gain access to a system for example by posing as staff to discover system vulnerabilities.

One common type of cybersecurity attack which makes use of computers is a DDoS which stands for a Distributed Denial of Service attack. This type of attack is difficult to guard against as it requires little specialist knowledge of computer hacking or the specific architecture attempting to be compromised. In a DDoS attack several connections are made to a website or other network resource and when too many connections are attempted to be made at once the system will be knocked offline which prevents non malicious users from accessing the resource. This can be achieved with standard hardware and coordination amongst large groups or through subversion via viruses. Although it is not easy to access sensitive information directly through this type of attack, it can be extremely damaging as the loss of revenue to a company can seriously effect its financial standing. Sophisticated networks of users can also mask their network traffic using network tools such as Tor, which also provide access to web services which are unreachable through a standard web browser. This technology was originally developed by the American Navy and has since found popularity amongst criminals but also journalists and whistle-blowers.

The effective sharing and storing of sensitive data can be conducted through many different types of encryption. RSA Encryption was one of the first attempts at creating an algorithm which suits the needs of digital security. It was developed in the United Kingdom approximately 40 years ago although has only been public knowledge for half of that time as it was classified information. The name RSA is taken from the surnames Rivest, Shamir and Adleman who were the first people to publicly describe the algorithm. Essentially the algorithm can be described with the following psuedocode:

1. Select two large different prime numbers at random

2. calculate the product of these two numbers

3. find the number of integers which are lesser in value to the original prime numbers chosen and also co-prime

4. for the value calculated in step 3, find a smaller integer which is also co prime to the number calculate in step 3. This is the public key

5. a congruent relation is calculated with the value from step 4. This becomes the private key

This algorithm allows the public key to be shared with anyone and used to encrypt traffic or files which, under currently known methods, cannot be decoded without knowledge of the original prime numbers chosen (assuming they are large enough). Similar systems are also used as one drawback of this algorithm is that it can take a long time to encrypt files however it is effective at sending messages as it provides not just encryption, but also verification that a message has been sent from a particular individual. This algorithm relies on the fact that factoring large numbers takes a long time, however with breakthroughs in computing such as quantum state computers it is possible that this system could be laid open overnight.

The most secure systems will make use of layers of technology and more complicated algorithms such as biometric security or ESA. An interesting point made by the speakers is that security breaches are treated as a certainty not a possibility. The importance of innovation was also made important.

Winlab (2008), Public Key Cryptography: RSA and Lots of Number Theory

Computer systems are generally developed to aid mankind in our endeavours. In recent years the challenges overcome and expected to be solved by and with computers is growing. To meet these expectations computer systems are constantly being developed with innovative ways of carrying out calculations. Using our own intelligence as a benchmark in chess ability or even normal conversation has been an aspect of artificial intelligence since the field was first conceived. With advances in hardware, huge strides are being made at a pace not previously imagined. In the last year, it was expected that the classic board game Go would be dominated by humans for at least another 5 to ten years however in the last few months there have been thinking machines developed which have triumphed over current world champions. The subject of the talk Given by IBM focused around one of the most well known AI systems: Watson.

Watson is essentially a set of Software API’s which can be added into existing apps. These include virtual assistant conversations which can provide customer service, dictation software to convert speech to text and even machine vision which is essential for the development of driverless vehicles. IBM also provides software products under the Watson product to aid in these fields and many others. This generates income for IBM but also has applications in academia outside the field of computer science. In a previous talk given by IBM on Watson at the University of Exeter, Watson was shown to have aided researches in cross referencing scientific papers at a rate that human scientists would be hard pressed to match to large success. Watson was able to identify a rare cancer and recommend treatment. This analysis of large datasets is an important aspect of thinking machines in general and is not to dissimilar to how we make decisions.

A good example of how Watson makes decisions is also what the system is most known for, which is besting some of the greatest human players at the quiz show Jeopardy. Watson was not connected to the internet and had to answer questions in the same time as its veteran human opponents. A simplification for how decisions are made is listed below.

1.anaylise the voice of the host for keywords.

2. select from the data set relevant answers to return

3. from the top three answers, assign each a confidence level

a. if the confidence level for all answers is under a set value, do not answer

b. otherwise, activate the contestant buzzer with the answer with the highest confidence

The technical specs of Watson while powerful, are not overpowered for a supercomputer. It is clearly the implementation of the software that allowed this machine to beat its human components at this task. Although winning game shows is not the stated goal for Watson or IBM, it generated good publicity for artificial intelligence in general and also resulted in software products that have since become commercially available. The same technique can be used in oncology treatment with slight modification in that instead of answering several short and relatively simple questions, it instead answers one complicated question with ore time to consider different answers.

Improvements to the Watson system are focused on the hardware side by taking advantage of cloud storage technology and on the software side by taking its question answering and language processing skills and applying them to different situations. Other commercial applications were explained during the presentation such as assisting customers within well known retailers, to offering cooking advice based on available ingredients.

Some interesting questions were posed to the speaker which included explaining the pros and cons f supervised learning as apposed to unsupervised learning. Having the human oversight at this stage is deemed more beneficial to the overall learning cycle although this raises the more philosophical point of who is then doing the learning.

IBM Research (2011) "The DeepQA Project".