

|  |
| --- |
| Advanced Research Computing  2020 Annual Report |
| Andrew Edmondson, Aslam Ghumra, Carol Sandys, Simon Thompson  January 2021 |

|  |  |  |  |
| --- | --- | --- | --- |
| Annual Report  2020 | | | |
|  | Introduction by Carol Sandys,Acting Head of Advanced Research Computing 2020 posed many challenges, not least for the teams providing the technology on which the University’s research depends. Despite the obstacles, both personal and practical, this has been a year of clear achievement for Advanced Research Computing (ARC) as the team forged ahead to deliver the Research Computing Strategy, responded to the COVID crisis, supported researchers to win funding and continued to maintain service as usual for an ever-growing constituency for whom computing has become fundamental.  Members of ARC have been on campus every week throughout the year, trouble shooting, fixing, upgrading, installing, commissioning and doing all of the essential engineering required to ensure the BEAR infrastructure has been in a position to support research in a “business as usual”. The rest of team has been operating efficiently from home, responding to user requests or problems and innovating to give research staff and students vital access to key software tools while off campus. COVID lockdown drove the accelerated delivery of the BEAR portal, providing access to common graphical software used often used on workstations. Together with Lenovo, ARC was able to specify, source and install £200k’s worth of equipment and adapt processes to support the sequencing and analysis of COVID samples as part of the COG-UK project. And, remarkably, do it within 2 weeks.  The Vice Chancellor’s Review of IT Services in 2015 recognised the value of BEAR and the need to invest further to unlock more benefit for the University’s researchers. The results of that investment are evident in a host of ways whether that is the recognition from our peers that Birmingham does “research computing right” and is the model to follow, or recognition from the Tech industry through awards in the face of global competition or through acknowledgement in academic papers for the essential contribution of a research software engineering support, or feedback from individuals who have learned new skills and unlocked new paths for their research.  Unlike many of our peers in the Russell Group, the goal at Birmingham for the provision of advanced computing tools for research has never been to be the biggest. Instead, the approach is to be the best; best in terms of quality, capability, value for money and the ability to meet the needs of researchers from across the full spectrum of disciplines and to do it with minimal delay. To that end, BEAR provides supercomputing power to both traditional users of High-Performance Computing in the physical sciences and wide-spread adoption across other disciplines; those working with financial models; life-sciences researchers doing life-saving research. Beyond classical supercomputing, BEAR also delivers a range of complementary services, from storage, through data transfer, AI or code development and testing. e.g. for those working with ‘Big data’ or working in the field of Computer Vision or with other applications of AI; all underpinned by our ‘data everywhere’ policy.  The creation of ARC’s Research Software Group (RSG) is one of the key advances of the past five years. ARC now has the capability to train and coach core software skills, to carry out software projects including several which are at the heart of REF impact case studies, e.g. everydaylookism.bham.ac.uk  This Annual Report illustrates some of these achievements and gives key statistics about BEAR and ARC.  You can easily change the formatting of selected text in the document text by choosing a look for the selected text from the Quick Styles gallery on the Home tab. You can also format text directly by using the other controls on the Home tab. Most controls offer a choice of using the look from the current theme or using a format that you specify directly. | | |
| Page 1 | | | |
| Direct Infrastructure Investment | | | |
| £5m Projected research group infrastructure spend from special projects. £130k Projected research group infrastructure spend from premium investment. 14 days Time from order to delivery of COG-UK storage platform leveraging vendor support. UK First The Baskerville system is expected to be the first UK deployment of the NVIDIA Redstone 4GPU board. 9509 Infrastructure data points monitored every few minutes to validate system health. | | Research Group Investment Research groups continue to invest heavily in the BEAR infrastructure and ~£5m in directly awarded research grants are expected to have been invested in the wider BEAR infrastructure in 2020. This income has mainly been sourced via two research areas:   * **Baskerville** – the national Tier 2 HPC facility, led by Professor Iain Styles which resulted in Birmingham receiving £4m in funding from EPSRC;   Awareness of the Tier 2 funding opportunities enabled ARC to support development of an accelerated compute platform, leveraging our technical partnerships and reputation of Birmingham’s supercomputing facilities. The Tier 2 facility will be the first such Birmingham-led project.   * **CLIMB** – led by Professor Nick Loman, the centre has been successful in obtaining ~£1m in MRC funding to develop the CLIMB platform and support;   Whilst some funding was expected to support development of CLIMB, significant additional funding was made available through COVID research programmes, including “*MRC World Class Labs*” funding, made possible by our exceptional support of CLIMB.  In addition to these high-profile research programmes, a number of researchers have made more modest investments in “premium” BEAR services, expected to be ~£130k in 2020:   * Dedicated access to compute, including a significant investment from Dr Andrew Morris for an external engagement research programme * Top-Up Storage to add to the default allocation of 3TB per project funded from the IT Services capital allocation (via the Estates charge on research grants). The increasing use of the Research Data Store has driven significant and much needed improvement in the management of research data and reduction in the institutional risk from a multiplicity of local and often unreliable storage solutions.   Premium storage is also awarded to projects by the CaStLeS Exec as strategic investment in life sciences research.  ARC routinely supports research bids both pre- and post- award to support researchers’ acquisition of funding to invest in infrastructure and storage to help support pump priming of projects in Life Sciences. | |
| Page 2 | | | |
| Research Software Engineering | | | |
| £45k Projected research group investment in Research Software Engineering. 292 New application installation requests for research software on the BEAR Infrastructure. 21 Coaching sessions. 191 Training attendees. 2000 Automated software tests running every night. | | Research Software Engineering ARC’s Research Software Engineers (RSEs are fundamental to the operation of BEAR, building and supporting the broad and rich range of software applications that researchers employ, supporting users to exploit the computational services and constructing the system that administers the user resources. In addition, they have a mission to improve the use of software in research activities and to participate in the wider national/international RSE community as a key component of the UK’s national e-infrastructure. Alongside building a community of coders, training researchers, this work includes short engagements providing advice, coaching and coding at no cost to the research group. As a consequence, the investment total for 2020 is modest though the list of achievements and plaudits proves the value of this core service. A few examples are listed below:   * Development of the UK Coronavirus Cancer Monitoring Project’s websites: <http://ukcoronaviruscancermonitoring.bham.ac.uk> * Coaching a PhD student to learn the necessary software skills to enable high quality data analysis in their studies * Web applications development to support a series of REF impact cases, including for Professors Heather Widdows and Michaela Mahlberg * Solving memory access errors in a researcher’s GPU (CUDA) code * Porting a research group’s MATLAB software to C++ to allow much wider international use by clinicians and researchers.   To see a fuller report on their activity with a series of case studies, see the RSG’s annual reports: <https://intranet.birmingham.ac.uk/rsg-reports>.  “*In the final phase of development [of our research software], the team has relied on the support from the [RSG]. This support resulted with crucial improvements in robustness, consistency and clarity of the code. Our experience in collaborating with the BEAR team is nothing but excellent. It is a testimony of both the importance of RSE, in general, and the quality of the BEAR support, more specifically. We are looking forward to our future collaboration on our new projects.*”  – Dr Petar Milin, Modern Languages, <https://outofourminds.bham.ac.uk/>”  In 2021, a new service will be added with the intention of offering extended RSE engagements for substantial coding tasks to be funded directly by research groups. We anticipate this will grow with the appointment of the first pool RSE. | |
| Page 3 | | | |
| Collaboration and Training | | | |
| Training COVID-19 restrictions limited our ability to deliver in-person training during 2020. With some development of the materials and substitution of instructors, we were able to deliver a number of courses on-line and are in a position to continue in the coming year. The departure of one member of ARC in December sees the loss of an NVIDIA Deep Learning Institute instructor which will affect the ability to deliver NVIDIA accredited courses for some time.  Training is offered to researchers, both staff and students, to enable them to build the foundation skills to use BEAR services and to learn how those services, including software engineering can improve their research both in terms of quality, scale and speed to results. In-house provision allows the course to be tailored to the needs at Birmingham and to reduce costs such as travel, subsistence and time, making it widely accessible. External Speaking Engagements  * Dr Simon Branford presented on ‘EasyBuild and POWER9’ at the EasyBuild User Meeting in Barcelona * Dr Simon Branford presented at UK HPC Champions about Baskerville, the new EPSRC funded Tier2 service for accelerated compute. * Dr Stephanie Thompson was invited to speak at the Research Data Alliance 16th Plenary Meeting - Title of the breakout session talk was ‘BEAR and Pizza - the route to engage researchers with data’ * Dr Stephanie Thompson gave a lightning talk at the IDCC20 in Dublin on ‘Providing Software Support to Enable Research: From Feral Parakeets to the Times Digital Archive’ * Simon Thompson continues to lead IBM’s Global Spectrum Scale User Group as Chair. Coordinated virtual events and spoke on several occasions * Simon Thompson spoke at CIUK as part of a session on the University’s approach to sustainable HPC  University Events and Meetings  * Regularly presenting to research units of varying size from College or School level to Research Group e.g. School of Maths, CHBC and the new CAL DoR. * Continuing partnership with Prof Jean-Baptiste Cazier and with the wider CaStLeS Exec * Growing partnership with Prof Iain Styles as the newly appointed Director of BIIDS&AI * Regular engagement with the Research Computing Management Committee (RCMC) with its representatives of all five Colleges * Supporting PVC Prof Tim Softley’s Academic IT Reference Group * BEAR Champions – a network of research group champions has been developed and fostered  National and International Representation:  * CLIMB-BIG-DATA Planning * Tier2 bid planning, liaison with Diamond Light Source, the Rosalind Franklin and Alan Turing Institutes * Regular meetings of the Russell Group Director’s Research IT Sub-group   Active participation in the national HPC Special Interest Group IT Services  * Member of the ITS Identity Steering Group * Member of DNS Steering Group * Supported ITS Go, including pre-COVID lockdown events * Teams pilot programme engagement | | |
| Page 4 | | | |
| Research Data Store | | | |
| 1.2B Over 1.2 billion files are hosted in the Research Data Store. 2.7Pb Over 2.7Pb of data are currently stored in the research data store. 3M Over 3 million files regularly change during the day and are processed by the backup system.  Chart, sunburst chart  Description automatically generated | | Research Data Store The Research Data Store (RDS) is made available for any research project, with up to 3TB of space available for each research project. Additional storage is made available to purchase at a competitive rate, specifically this is to encourage the use of the facility over other approaches such as local USB hard-disks. All data held on the RDS are backed up with a second copy held on tape in a separate building.  Chart, line chart  Description automatically generated  The graph shows the continued growth in research data held and safeguarded in the RDS. A good proportion of the growth represents take-up by additional research groups as well as the increased intensity/scale of research data particularly in the Life Sciences. In many cases, growth in the RDS is matched by reduced demand for local storage in Colleges/research groups. The RDS caters not only for users of our supercomputing services but also for all researchers who work with valuable data and collaborate within and external to the University. It serves many of Birmingham’s top-earning researchers as well as some of our most highly cited.  The final portion of the graph represents the replacement of the ‘end of life’ CaStLeS storage pool, and reclassification of some historical RDS projects as CaStLeS. The programme of works moved all the data into a single name space and was achieved in collaboration with research groups to minimize inconvenience by scheduling and migrating data on a project basis. In data management terms this was a massive exercise with circa 360 million files moved.  The new CaStLeS storage pool adds capacity. At the same time, the team updated some of the aging underpinning infrastructure to improve performance and resilience.  Daily change and ingest rates have proved a challenge to our ability to effectively backup the data, a significant overhaul of the backup system had been planned for 2021, however this has been scaled back to allow spending to be deferred. 2021 will however still see changes made to the backup infrastructure to ensure data is retained safely. | |
| Page 5 | | | |
| Compute platform | | | |
|  | | | |
| 7046 Peak concurrent cores in use – Dec 2020. 902 Registered BEAR Compute Projects. 3 Architectures supported – Intel “x86”, IBM POWER9, GPU accelerated. 100Gb Low latency interconnect provides 100Gbps speeds to accelerate multi-node and data intensive workloads. | | Supercomputing BEAR’s compute services are versatile, built to serve the breadth of the University’s computationally or data intensive research. Representing this simply is a challenge but the graph above, gives an idea of the continuing growth in demand for compute and the heterogeneous nature of the core service.  In 2020 we saw a new peak in terms of the point in time demand for compute power that we were able to meet. Continuous sustained load is present throughout the year, and lockdowns due to COVID-19 have had little impact on the demand of the system.  Along with a cohort of big consumers and funders of the service, BlueBEAR supports over 900 projects which queue for their turn when the specific resources requested become available. At present typical wait times are healthy and this is the balance that must be maintained. Each year a generation of compute nodes becomes obsolete and is retired to be replaced with a tranche of new equipment which is not only faster and more capable but, typically, much more energy efficient and sometimes based on radically new technology. This generational churn allows us to continue to meet growing demand for both scale and performance but to do so on the basis of steady state investment. Nevertheless, growth in the number of researchers and research projects with a fundamental need for this service will accelerate demand and potentially increase the investment required to maintain service standards.  In April 2020, the BEAR portal was launched. This provides a web-based mechanism for accessing the batch facilities and includes support for graphical applications. Development of this was accelerated due to COVID-19 as a mechanism to provide access to graphical applications often found on inaccessible lab workstations. E.g. ANSYS  Beyond BlueBEAR, BEAR Cloud provides specialist virtual machine capability for workloads not suited to batch compute. For example supporting Galaxy pipelines and data intensive web-sites. | |
| Page 6 | | | |
| 2021 Outlook | | | |
| 2021 promises to be a busy year despite the demands and limitations imposed by COVID. As always, ARC’s programme remains dynamic to ensure it is able to respond to the needs of new research groups and projects as well as sustain service levels for all users of BEAR. Our aim is to ensure the service remains amongst the best serving a Russell Group institution and is demonstrably value for money.  Plans have been revised in response to COVID budget constraints, resulting in the deferral of some activity. Nonetheless ARC has an ambitious schedule including activity in a number of key areas.  It has become increasingly evident that there is a desire to improve reporting metrics on the various aspects of the service, to support capture of additional funding, and to report how research outputs are supported by the service. During 2021, ARC will use the annual BEAR project re-registration process to collect additional data related to each project, e.g. WorkTribe project number, whilst being flexible to recognise that not all research outputs are funded. ARC will also continue to work with Corporate Services partners to support wider inclusion of BEAR premium services in funding bids, though it is important to note that the research support teams only see a small proportion of grant applications. To truly deliver impact, the University needs to invest in changes to the grant application and reporting tooling, for example capturing requirements in WorkTribe and outputs in Pure, better support for internal charging and acknowledgement that income may need to be held across financial years is important, but outside of the scope for ARC to make changes.  Similarly, with access to new Tier 2 resources, this presents an opportunity to streamline and consolidate the resource allocation process and discussions are ongoing on how a revised “academic allocation board” may be constituted.  The implementation and launch of Baskerville will take significant resource from the team to implement an efficient service and aims to launch for EPSRC RAP respondents in July 2021. Whilst the award of the grant is beneficial to the University, offering a poor service level could adversely affect the University’s reputation. It is hoped that the two Birmingham based posts will be in-post in time for the launch.  Whilst it is unclear how long COVID-19 will continue to impact general operations, the ongoing support of CLIMB and the COG-UK work will continue for some time and the impact of the success of these projects is likely to have long-term service and funding implications beyond the current offering.  Recruitment of RSEs to fill both a vacancy and new post to support the demand for funded longer term engagements is underway and this will support the pool RSE service. Several funded engagements are already in the pipeline for 2021.  During 2020, time was spent with Prof Iain Styles in the development of the BIIDS proposal, in addition to this, requirements from Prof Jean-Baptiste Cazier to support limited teaching was necessary. Whilst the BEAR environment is not resourced to support teaching, we expect this requirement and additional requirements from BIIDS to widen the scope. To support this effectively in 2021, additional resource will be required within ARC.  External factors will continue to provide significant challenge to the development of the facilities. COVID-19 continues to have an impact on both our ability to work effectively in the data centre spaces and our ability to safely employ external engineers when required. The impact of BREXIT will also likely have impact, in the form of delays to shipments or increased costs. “Next-business-day” parts for hardware faults have been held at the border by HM Customs and Excise and this raises risks to operational service.  Changes to the hardware roadmaps of manufacturers may also result in significant disruption to plans with anticipated delays to the development of the next generation of CPU. Switching vendor would pose non-trivial integration challenges and likely industry consolidation will introduce uncertainty to this complex market. Close working with vendors will be required to support the annual renewal programme of hardware.  Notwithstanding all this, the ARC team are resilient and dedicated to providing excellent service to research and look forwarded to the new challenges of 2021. | | |
| Page 7 | | | |