REF2021

Institution: University of Glasgow (UofG)	

Unit of Assessment: UoA 11 Computer Science and Informatics

Title of case study: Increased scalability of Erlang for improved telecommunications and internet server capability.

Period when the underpinning research was undertaken: April 2011–June 2017

Details of staff conducting the underpinning research from the submitting unit:

Names(s):	**Role(s) (e.g. job title)**	**Period(s) employed by the submitting HEI:**
(1) Phil Trinder	(1) Professor, Computing Science	(1) April 2013–present
(2) Natalia Chechina	(2) Research Assistant	(2) June 2013–June 2017

Period when the claimed impact occurred: December 2012–present

Is this case study continued from a case study submitted in 2014? No

1. Summary of the impact

Erlang is an open-source specialist programming language governing >90% of internet traffic and ~50% of landline phone calls globally. Trinder at UofG identified fundamental issues which constrained Erlang's performance on multiple cores (processors) and hosts (machines), i.e. its scalability. Trinder initiated, assembled and led an international academic/industrial consortium to address these issues. UofG research designed, developed and evaluated better language-level scalability for Erlang, creating the Scalable Distributed (SD) Erlang libraries. These allow Erlang to deliver fault-tolerant, scalable programmes (by organisations including Ericsson, WhatsApp and NHS Digital) for tens of thousands of applications in areas such as fintech, security, blockchain and Internet of Things.

2. Underpinning research

Erlang is one of the world's most widely used programming languages, designed to produce scalable, fault-tolerant, concurrent, distributed, non-stop, soft real-time applications. Around 2009, UofG's Prof. Trinder analysed the scaling issues associated with Erlang, which made it difficult to obtain good software performance on large multi-core computer architectures, e.g. servers. Prof. Trinder initiated, established and led, as EU project coordinator, an international academic consortium (RELEASE http://www.release-project.eu/) to overcome these issues. The RELEASE project focused on addressing the increasingly critical scaling problems with Erlang, and hence addressed the widening gap between state of the art in hardware and software.

While leading the RELEASE consortium, UofG's Dr Chechina and Prof. Trinder designed, developed and evaluated better language-level scalability for Erlang, in the form of a new library: Scalable Distributed (SD) Erlang. The new technology is employed on servers and in conjunction with the Erlang *Open Telecom Platform* (OTP).

RELEASE tackled these scalability challenges at three levels:

1. Developing infrastructure to deploy and measure the performance of Erlang on such large-scale architectures
2. Developing and validating the Scalable Distributed (SD) Erlang libraries that enable Erlang to be deployed on large scale architectures, i.e. clusters of multicore machines
3. Evolving the Erlang Virtual Machine (VM) so that it can work effectively on large multicore machines.

Scalable Erlang Infrastructure	
Scalable Distributed (SD) Erlang Language	Tools
Erlang Virtual Machine	

The unique UofG contributions to the RELEASE consortium led by Trinder were:

- Prof. Trinder's identification of the fundamental scalability issues with Erlang in 2011, specifically at the Virtual Machine, Language, and Infrastructure levels.
- Prof. Trinder designed solutions and ways of tackling the scalability issues with key industry and academic players including Ericsson, Erlang Solutions and Kent University. This resulted in the foundation for how the new form of Erlang could be developed [3.1, 3.2].
- The vision paper that described the successful EU FP7 research proposal. There were three primary authors Phil Trinder (Glasgow), Simon Thomson (Kent), Ulf Wiger (Erlang Solutions). The paper describes the proposed research. Natalia Chechina, a Glasgow PDRA on the RELEASE project, is the lead author (currently in Bournemouth).

Following the development of SD Erlang, a UofG Knowledge Exchange (GKE) grant was awarded to Trinder (2015–2017) to maintain the cyber presence of SD Erlang and RELEASE, e.g. website, ensuring it contained up to date information and could maximally benefit the programming community [3.3]. This funding also allowed the maintenance and restructuring of the SD Erlang libraries. After discussions with industrial users of SD Erlang, it was noted that in the dynamic world of Erlang/OTP development with yearly major releases and quarterly minor releases, the SD Erlang users preferred up-to-date versions of SD Erlang to use in production, rather than well-tested but outdated versions. Therefore, the UofG team automated the rebasing of SD Erlang for new releases of Erlang/OTP. The primary activities undertaken to achieve this were:

(i) Rebasing SD Erlang manually twice to versions 19.2 and 19.3. The automated rebasement of SD Erlang will be available from the major Erlang/OTP release 20.

(ii) Restructuring SD Erlang, separating some specific SD Erlang functions from global.erl.

Prof. Trinder's leadership of the RELEASE consortium has led to numerous key achievements including the development of WombatOAM, a proprietary software tool providing a scalable infrastructure for the deployment of Erlang/SD Erlang to thousands of computers.

3. References to the research

3.1 Boudeville O. Cesarini F. Chechina N. Lundin K. Papaspyrou N. Sagonas K. Thompson S. Trinder P. Wiger U. RELEASE: A High-level Paradigm for Reliable Large-scale Server Software. The 13th Symposium on Trends in Functional Programming (TFP'12), St Andrews University, UK (June 2012). Springer Verlag LNCS.

This paper appeared at the start of the RELEASE project and outlined the FP7 Research proposal and plan.

3.2 Trinder P. Chechina N. Papaspyrou N. Sagonas K. Thompson SJ. *et al.* Scaling Reliably: Improving the Scalability of the Erlang Distributed Actor Platform ACM Transactions on Programming Languages and Systems (TOPLAS) V39(4) (September 2017) pp 1–46. doi: 10.1145/3107937

This paper appeared at the end of the project in the premier Programming Language journal. It summarises the challenges addressed by the RELEASE project, and the progress made.

3.3 Chechina N. Huiqing L. Ghaffari A. Thompson S. Trinder P.W. (2016) Improving the network scalability of Erlang. Journal of Parallel and Distributed Computing, *Elsevier* V90–91 pp 22–34. (doi:10.1016/j.jpdc.2016.01.002)

This paper appeared mid-project in a premier distributed systems journal. It presents the rationale, design, semantics, implementation and validation of the SD Erlang libraries.

3.4 Chechina N. MacKenzie K. Thompson S.J. Trinder P.W. Boudeville O. Fordos V. Hoch C. Ghaffari A. Moro Hernandez M. (2017) Evaluating Scalable Distributed Erlang for Scalability and Reliability. IEEE Transactions on Parallel and Distributed Systems (TPDS). (doi: 10.1109/TPDS.2017.2654246)

Appearing in a premier distributed systems journal, this paper evaluates the effectiveness of SD Erlang.

3.5 Chechina N. Moro Hernandez M. Trinder P.W. A Scalable Reliable Instant Messenger using the SD Erlang Libraries. Proceedings of the 15th ACM SIGPLAN Workshop on Erlang, Nara, Japan, pp. 33--41 (September 2016) ISBN 9781450344319. (doi:10.1145/2975969.2975973) This paper demonstrates a realistic application using SD Erlang.

Grants

Phil Trinder (Coordinator); European Union Seventh Framework Programme (STREP grant 287510); EUR2.4 million (Oct 2011–Feb 2015 [Apr 2013–Feb 2015 at UoG]): RELEASE: A High-Level Paradigm for Reliable Large-Scale Server Software.

Phil Trinder; Glasgow University Knowledge Exchange Fund, GBP50,000 (Oct 2016–Jun 2017): impact acceleration project to enhance impact generation of new SD Erlang capacity.

4. Details of the impact

Today's world demands highly *scalable* computing: billions of people are simultaneously online for work and leisure activities; millions of companies are utilising software systems; millions of financial transactions flow through global systems. Computation must be correct, fast and seamless to satisfy demand and remain competitive.

Technological innovations emerging from UofG and through the RELEASE project delivered impacts on commerce and economy through the enhanced scalability of Erlang. The RELEASE project delivered improvements to Erlang/OTP from R15B03. Prior to the UofG input, a critical constraint on Erlang had been the lack of scalability, which increased the complexity of engineering scalable applications such as NHS Digital and WhatsApp.

The open-source Erlang language has users in major sectors including telecommunications, banking, online gaming and social networks. The broader economic impacts are diverse and, in some cases, hard to quantify. However, with over 40% of the world's mobile traffic and over 50% of landline calls carried over Ericsson networks using Erlang, and 90% of all internet traffic going through routers and switches controlled by Erlang, the customer base served spans 180 countries across the globe. Maintaining these networks and ensuring they function optimally is critical to Ericsson's economic success and that of the industries and people using Ericsson mobile networks. Ericsson confirms that RELEASE:

> *'made 10 important changes and architectural improvements….These changes mean that Erlang is more responsive and works better on NUMA [non-uniform memory access] machines. Such improvements benefit almost all Erlang users and the current download rate is approximately 50,000 per month'* [5.1].

As part of RELEASE, Trinder and Chechina designed, developed and evaluated better language-level scalability in the Scalable Distributed (SD) Erlang library, with first modifications released in December 2014 [5.2]. Overall, the project has produced, or contributed to, 8 new open-source software tools which enhance the scalability and function of apps and communication services around the world [5.3–5.5].

The Erlang improvements made in RELEASE are now part of the Erlang/OTP software platform [5.6] that drives mobile communication and some of the world's most famous large-scale apps. The improved Erlang/OTP has enabled:

1. Improved capacity to work at a larger scale – it is far easier to engineer software for hundreds and thousands of hosts in SD Erlang than in Erlang; and
2. Reduced software development times, reducing time to market in a plethora of extremely competitive industry sectors.

Erlang's improved programming language technology is utilised by thousands of companies producing a wide diversity of applications. It has resulted in the production of successful apps [5.7a] in sectors as diverse as: communication (WhatsApp: 2 billion global Monthly Active Users [MAU]); gaming (Nintendo Switch); banking (Goldman Sachs, Mastercard) and personal finance (Klarna, 460,000 MAU in the UK alone); social networking (Grindr (3.8 million Daily Active Users [DAU]); real-time advertising (Adroll, Tapklik [5.8]); bookmaking (Bet365); web servicing (Amazon Web Solutions [5.7b]; and national providers of information data and IT systems (NHS Digital).

WhatsApp increased its user base by offering free international telecommunication irrespective of distance, platform and number of users. With 500 million daily active users, it is now one of the world's most popular apps on both Android and iOS platforms. High levels of scalability are required to simultaneously serve these users. Prior to RELEASE, WhatsApp had intricately engineered a scalable Erlang system. Now not only do they benefit from the RELEASE Erlang Virtual Machine enhancements, but achieve scalability using a re-implementation (with acknowledgement) of the core idea of SD Erlang as "Scalable Process Groups" [5.9].

The NHS Spine II system connects clinicians and patients to essential national healthcare services across the UK, providing the Electronic Prescription Service, Summary Care Record, the e-Referrals Service etc. It is used by 500,000 healthcare professionals daily and supports 28,000 IT healthcare systems across 21,000 organisations. It provides the central data storage and messaging systems for the NHS and uses two Erlang products: Riak and RabbitMQ. Riak is a leading internet-scale NoSQL database management system, while RabbitMQ provides a stable, well used, mature, open source, Erlang, queue implementation [5.10c].

Riak and RabbitMQ were improved by the RELEASE project and have been central to enabling Spine II to scale to meet rising user demand. Spine II handles over 10 million transactions a day (with up to 600 transactions per second). October 2020 was their busiest month ever, with more than 1.2 billion messages processed. This has been achieved with 100% availability on key messaging services. The enhanced response times of the new Spine II save the NHS 800 working-hours per day compared to the out-sourced, Oracle-based Spine I. In the first 5 years of its operation Spine II has saved the service GBP150 million, as its open-source design negates the need for expensive licenses for bespoke hardware and software [5.10a,b].

> "*Spine is the backbone of the NHS Digital infrastructure. The transfer of Spine II to a system based in Erlang has been a success story. These products combine fault tolerant resilient behaviour with the scalability necessary to facilitate health and social care organisations and workers access patient records when they need, wherever they need, to improve patient care*" [5.10a].

Erlang is also employed by Amazon servers in their business venture, Amazon Web Solutions (AWS), that allows users to build *'sophisticated applications with increased flexibility, scalability and reliability'* [5.7b]. Maximising their use of Erlang's scalability, AWS have dominated the market with a 34% share, well ahead of Microsoft at 11% (2017).

Bet365, a UK-based bookmaker, credited the switch from Java to Erlang for their ability to deliver innovative, competitive technology to their digital customers while also rapidly expanding their

customer base (from 2 million to 35 million customers across the globe) without increasing the risk of failure [5.11].

UofG continues to maintain the SD Erlang libraries [5.12] to ensure that worldwide users of those apps employing Erlang on their servers continue to benefit from SD Erlang's enhanced scalability. UofG expertise is also supporting industrial users of SD Erlang, such as TapKlik; this continued support has helped the company to gain investment due to the size of its international userbase, achieved as a result of the SD Erlang scalability enhancements. TapKlik [5.3] confirms that

> *'the standard Erlang release has a certain limitation in the way it organizes communication between nodes….SD Erlang…solves these issues'* and that
> SD Erlang is *'enabling Tapklik bidder to achieve better performance on the same resources and better scalability.'*

UofG continues to collaborate with the Ericsson Erlang/OTP team to assist them in applying scaling techniques at the Virtual Machine Level [5.1].

5. Sources to corroborate the impact

5.1 Letter from the head of Erlang/OTP team at Ericsson

5.2 RELEASE Exploitation plan: release-project.softlab.ntua.gr/documents/D7.3.pdf

5.3 Letter from Erlang Solutions identifying contribution to WombatOAM product

5.4 Webcasts/interviews at industry conferences:
   a. Erlang Factory Lite 2014 (San Francisco, USA) youtu.be/XncdY2KTfis
   b. Norwegian Developers Conference (NDC) 2014 (Oslo, Norway) vimeo.com/97537022
   c. Erlang User Conference 2014 (Stockholm, Sweden) youtu.be/dWpsesw_UQU
   d. InfoQ interview www.infoq.com/interviews/chechina-sderlang
   e. QCon 2015 (New York, USA) www.infoq.com/presentations/actor-functional-scale?utm_source=infoq&utm_medium=QCon_EarlyAccessVideos&utm_campaign=QConNewYork2015
   f. BuildStuff 2016 (Vilnius, Lithuania) youtu.be/G3cJI882wuQ

5.5 Github links for the 8 open source software tools can be linked from a summary web page: release-project.softlab.ntua.gr/index.php/software
   Concuerror (12 contributors); Devo (1); Dialyzer (20); ErLLVM (3); Percept2 (6); SD Erlang (3); SD-Mon (2); Wrangler (2).

5.6 Erlang/OTP website acknowledgement of the contribution of the RELEASE project: www.erlang.org/community under "Research Projects"

5.7 a) Blogpost, Erlang Solutions. www.erlang-solutions.com/blog/which-companies-are-using-erlang-and-why-mytopdogstatus.html
   b) Amazon Web Services homepage https://aws.amazon.com/

5.8 Letter from TapKlik identifying the value of SD Erlang to their company [PDF]

5.9 WhatsApp's reimplementation of SD Erlang's "Scalable Groups" as "Scalable Process Groups": github.com/max-au/spg/blob/master/README.md
   WhatsApp's acknowledgement of SD Erlang's "Scalable Groups" in "Scalable Process Groups" design www.youtube.com/watch?v=FJQyv26tFZ8 (minutes 7-8)
   www.youtube.com/watch?v=216NV-odxnE (minutes 17-19)

5.10 a) Testimonial from NHS Digital; Evidence related to NHS Digital: b) Transcript of interview with Principal Systems Engineer, Spine Core, NHS Digital; c) github.com/russelldb/rabl;

5.11 Evidence related to online gambling industry: a) sbcnews.co.uk/features/2017/05/30/erlang-solutions-functional-programming-bookies/ ; b) Erlang in Gambling/Online Betting (2014); www.youtube.com/watch?v=X7ZDqhlfru4 (minutes 37-38)

5.12 Current version of SD Erlang libraries developed and maintained at Glasgow github.com/release-project/otp/tree/19.3-rebased