Collaboration, Capabilities, and Impact in NZ eResearch: Bridging the Gap

Rob Elshire

Some things about me

1987	Unix on the mainframe
2001	Founding Director Illinois Genetic Marker Center
2007	Solexa Sequencing at Cornell
2009	Maize HapMap -- Science
2011	Genotyping-by-Sequencing -- PLoSOne
2011	GBS Service @ Cornell
2012	Project Management / Software Dev. / Bioinformatics
2013	Shift to NZ / NZinc approach / eResearch 2020
2014	NZ OSS Peoples Choice Award
2015	MBIE: Breeders Meeting & eResearch Refresh

Collaboration

Capabilities

Impact

A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it.

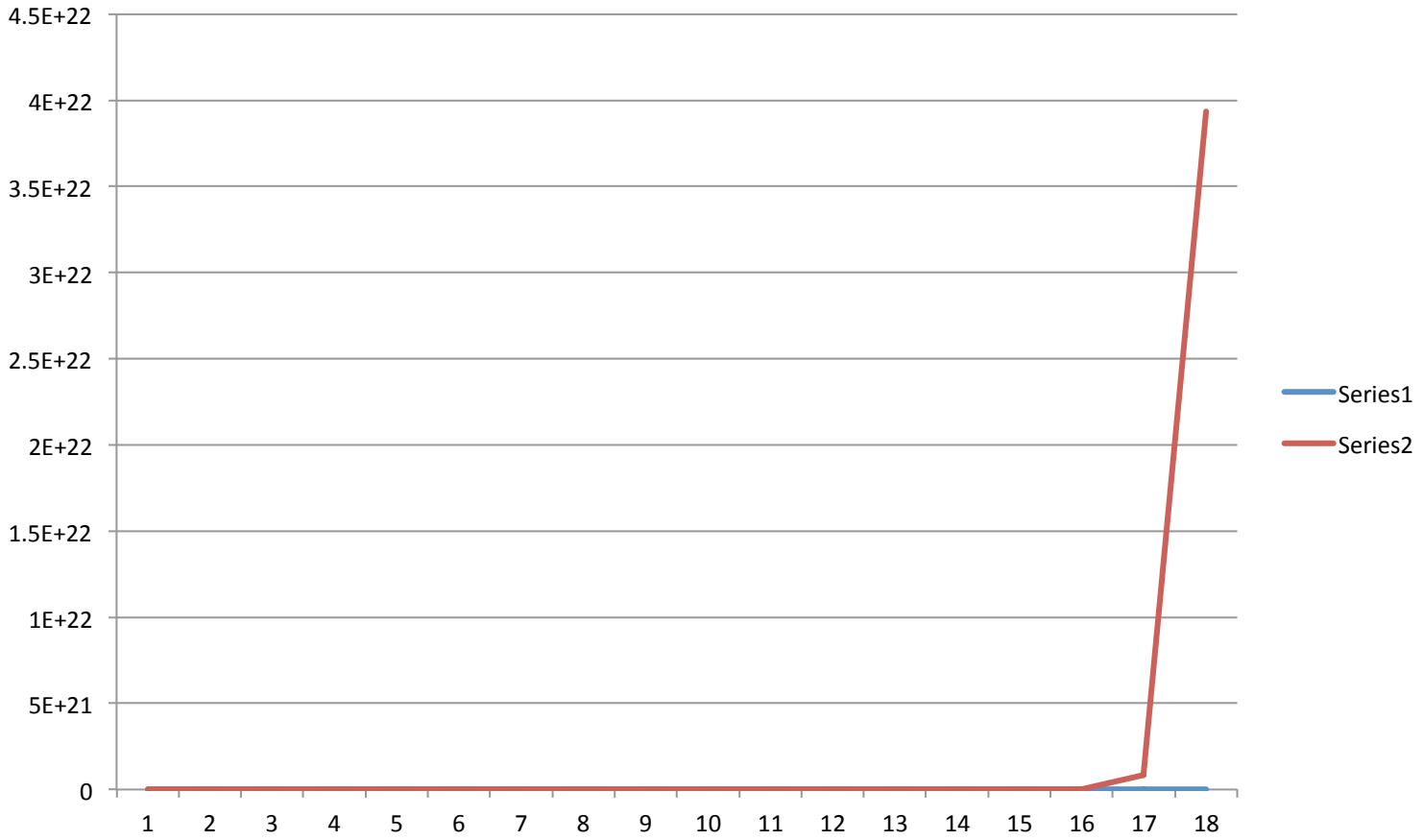
--Max Planck

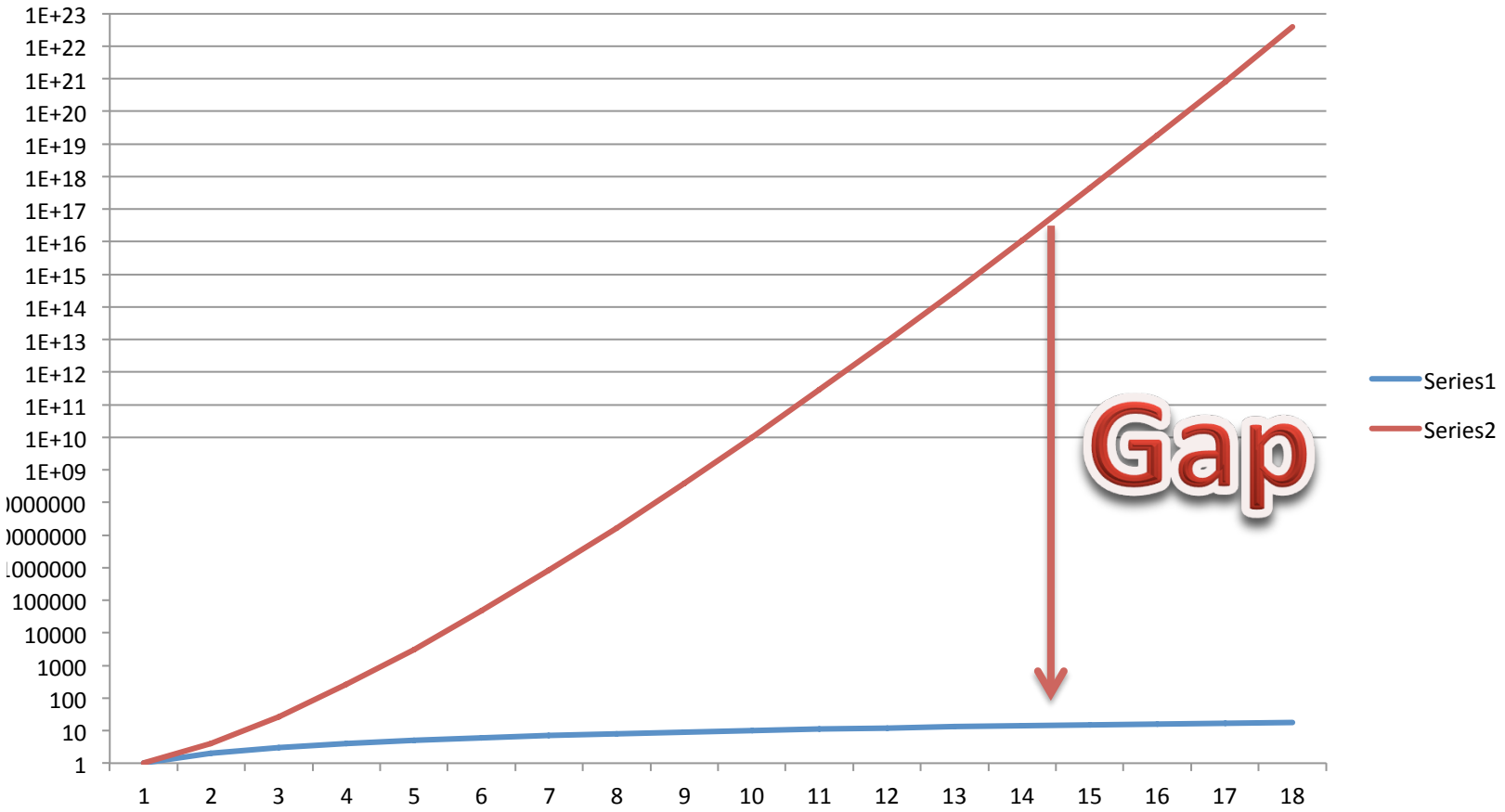
Gaps

Exponential vs. Linear Change

Exponential Growth

- Was noted by Inder Monga in the first plenary in terms of research networking.
- Even greater growth has been taking place in genomics research since 2007.
- Perspective & Resolution


4.5E+22
4E+22
3.5E+22
3E+22
2.5E+22
2E+22
1.5E+22
1E+22
5E+21
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Series1
Series2

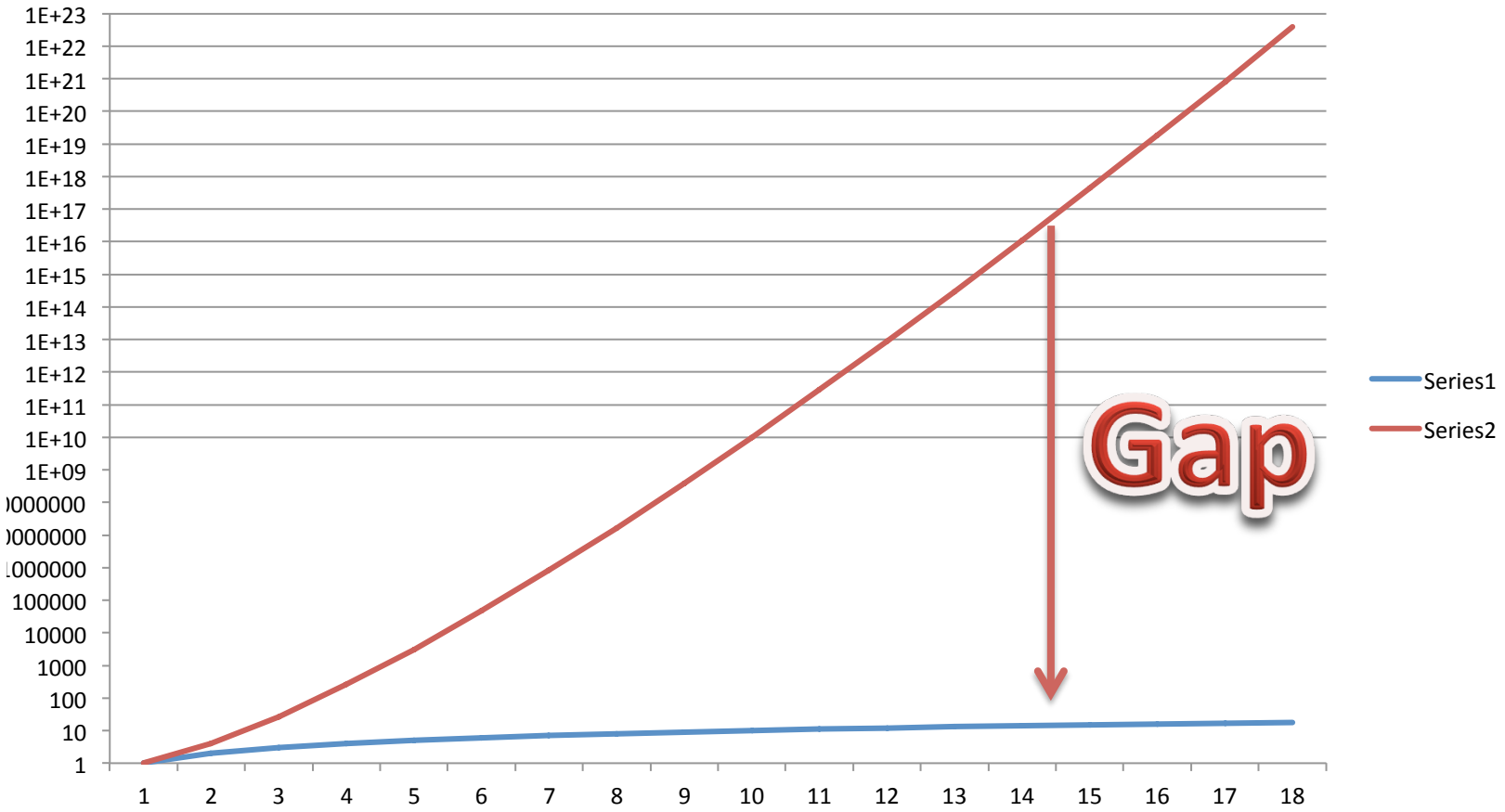

1E+23
1E+22
1E+21
1E+20
1E+19
1E+18
1E+17
1E+16
1E+15
1E+14
1E+13
1E+12
1E+11
1E+10
1E+09
)000000
)000000
1000000
100000
10000
1000
100
10
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Gap
Series1
Series2

Gap Year

- Many young people take a gap year between college and university.
- This time can be a period of personal growth.
- Global change in science continues to move at an exponential rate.

OE/Post-Grad/Sabbatical

- Joining the exponential global change
- What happens when they return?
- What about recruits from abroad?


1E+23
1E+22
1E+21
1E+20
1E+19
1E+18
1E+17
1E+16
1E+15
1E+14
1E+13
1E+12
1E+11
1E+10
1E+09
)000000
)000000
1000000
100000
10000
1000
100
10
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Gap
Series1
Series2

An International Perspective on the New Zealand Productivity Paradox

New Zealand Productivity Commission
Working Paper 2014/01

April 2014

http://www.productivity.govt.nz/sites/default/files/international-perspective-working-paper.pdf

Selected Points

- There is a 30% differential between actual productivity and expected productivity in NZ. This is the paradox.
- One component that contributes is a relative underinvestment in knowledge based capital.

eResearch Challenges in NZ Discussion Document

http://www.eresearch2020.org.nz/

SKILLS LAG: As a science system, we are under-investing in the broad-based research skills and methodological training that underpins the international shift in research methods to digital evidence and data intensive discovery. In addition, our major capital and programme investments are disconnected, either geographically, or organisationally. Consequently, we are putting at risk the quality of our research outputs — our stock of knowledge and our opportunity to grow high tech capabilities in the country, today and into the future.

RESEARCH COMMUNITIES: Impactful research collaboration usually occurs at the inter-personal, research-discipline level, not at project or institutional level. Greater explicit investment support for research-discipline led collaboration (rather than institutional collaboration), such as CoREs and NSCs, and support for discipline-based national research societies is likely to led to greater cohesion, knowledge exchange, and collaborative research outcomes across the system.

INCENTIVES DRIFT: Institutional funds, policies and practices for ICT are applied to research methods with potentially ruinous consequences, distorting research design, subverting resources, blocking collaboration, and leading to sub-optimal application of the balance of government investment in science overall.

FUTURE INFRASTRUCTURE: Major innovation or discovery in the coming decade is likely to stem from teams of researchers working across disciplines, institutions, and national borders, based on data from a proliferation of sources — including collaboration with industry — and with significant reliance on compute, integrated systems, sensor networks, and complex data analytics. Rather than lag behind comparator countries in science, provision should be made early for investment in capacity and capability in data, visualisation, and digital research expertise. New Zealand's science system might then continue to contribute to economic productivity and competitiveness, and improve population health, societal and environmental outcomes at first world levels.

Comparisons

- Cornell University
  - Sequencing Capacity
  - Compute Capacity
  - Bioinformatics
  - Programming
- University of X NZ
  - NZGL
  - REANNZ
  - NeSI

Business Computing / Scientific Computing

Right Tool?

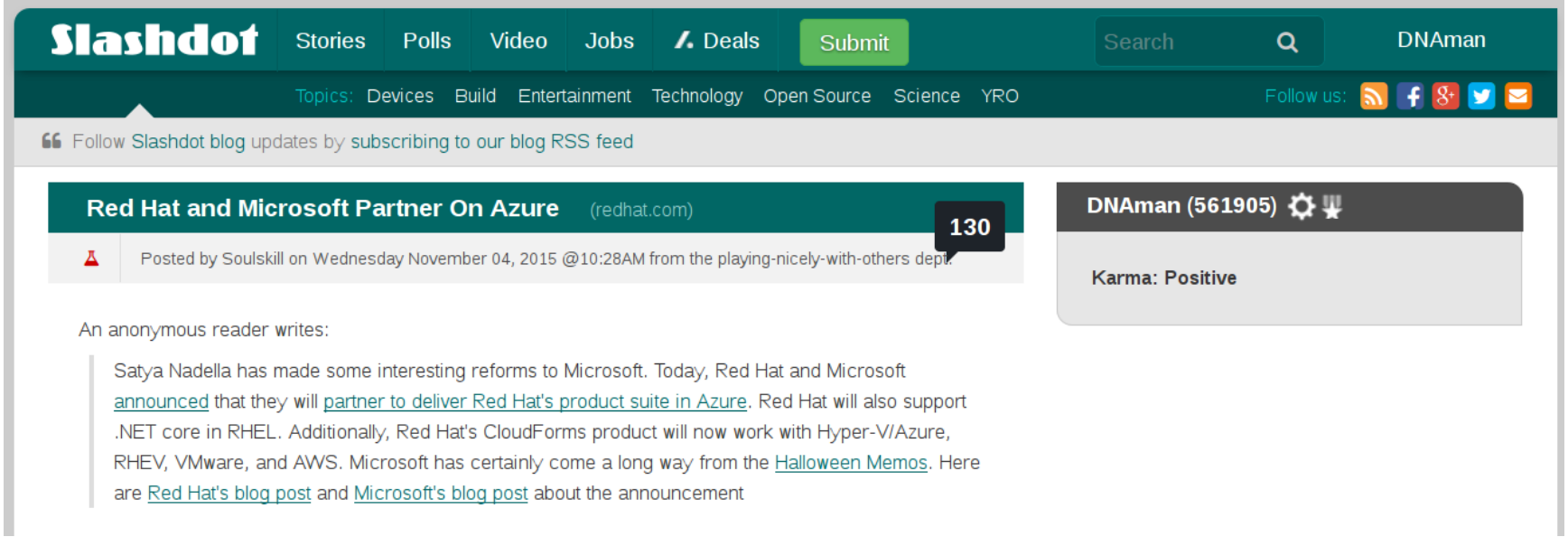
Microsoft Has Built a Linux Distro

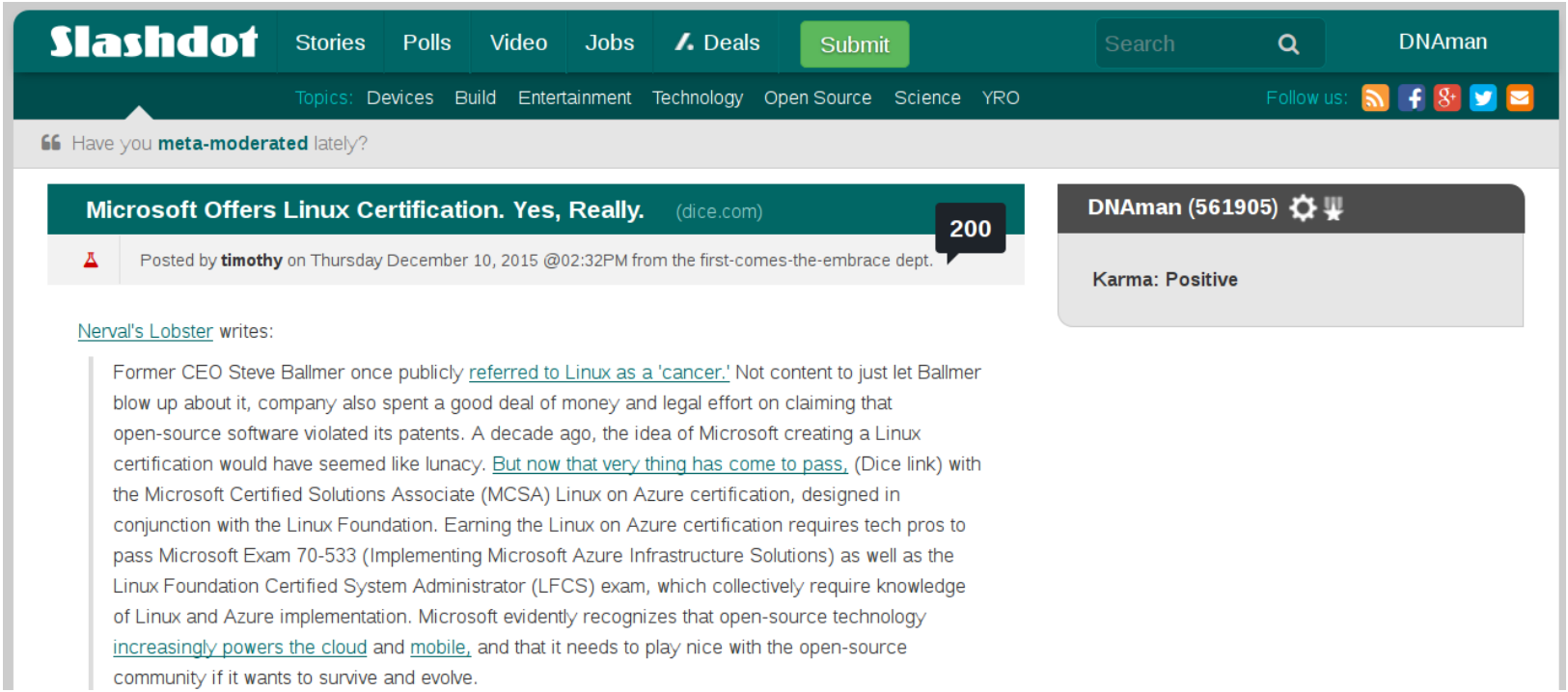
282

Posted by Soulskill on Friday September 18, 2015 @07:56AM from the wait-what dept.

jbernardo writes:

> Microsoft has built a Linux distro, and is using it for their Azure data centers. From their blog post: "It is a cross-platform modular operating system for data center networking built on Linux." Apparently, the existing SDN (Software Defined Network) implementations didn't fit Microsoft's plans for the ACS (Azure Cloud Switch), so they decided to roll their own infrastructure. No explanation why they settled on Linux, though — could it be that there is no Windows variant that would fit the bill? In other news, Lucifer has been heard complaining of the sudden cold.


Slashdot
Stories
Polls
Video
Jobs
Deals
Submit
Search
DNAman
Topics: Devices Build Entertainment Technology Open Source Science YRO
Follow us:
Follow Slashdot blog updates by subscribing to our blog RSS feed
Red Hat and Microsoft Partner On Azure (redhat.com)
130
Posted by Soulskill on Wednesday November 04, 2015 @10:28AM from the playing-nicely-with-others dept.
An anonymous reader writes:
Satya Nadella has made some interesting reforms to Microsoft. Today, Red Hat and Microsoft announced that they will partner to deliver Red Hat's product suite in Azure. Red Hat will also support .NET core in RHEL. Additionally, Red Hat's CloudForms product will now work with Hyper-V/Azure, RHEV, VMware, and AWS. Microsoft has certainly come a long way from the Halloween Memos. Here are Red Hat's blog post and Microsoft's blog post about the announcement
DNAman (561905)
Karma: Positive


Slashdot
Stories
Polls
Video
Jobs
Deals
Submit
Search
DNAman
Topics: Devices Build Entertainment Technology Open Source Science YRO
Follow us:
Have you meta-moderated lately?
Microsoft Offers Linux Certification. Yes, Really. (dice.com)
200
Posted by timothy on Thursday December 10, 2015 @02:32PM from the first-comes-the-embrace dept.
Nerval's Lobster writes:
Former CEO Steve Ballmer once publicly referred to Linux as a 'cancer.' Not content to just let Ballmer blow up about it, company also spent a good deal of money and legal effort on claiming that open-source software violated its patents. A decade ago, the idea of Microsoft creating a Linux certification would have seemed like lunacy. But now that very thing has come to pass, (Dice link) with the Microsoft Certified Solutions Associate (MCSA) Linux on Azure certification, designed in conjunction with the Linux Foundation. Earning the Linux on Azure certification requires tech pros to pass Microsoft Exam 70-533 (Implementing Microsoft Azure Infrastructure Solutions) as well as the Linux Foundation Certified System Administrator (LFCS) exam, which collectively require knowledge of Linux and Azure implementation. Microsoft evidently recognizes that open-source technology increasingly powers the cloud and mobile, and that it needs to play nice with the open-source community if it wants to survive and evolve.
DNAman (561905)
Karma: Positive

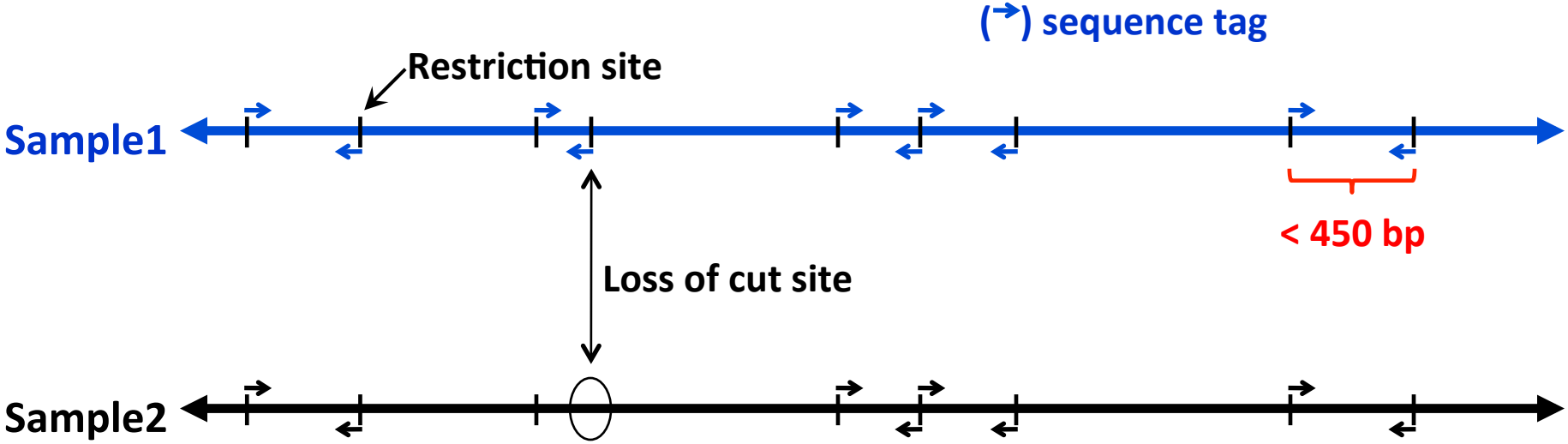
Genomics as an Example

Genotyping-by-Sequencing

What is GBS?

Genotyping by sequencing is a genotyping method based on restriction digestion and high levels of multiplexing which results in a library ready to sequence on the Illumina platform. Key features are simplicity, reproducability, speed and low cost.

Overview of Genotyping by Sequencing (GBS)



- Focuses NextGen sequencing power to ends of restriction fragments

Molecular Tools

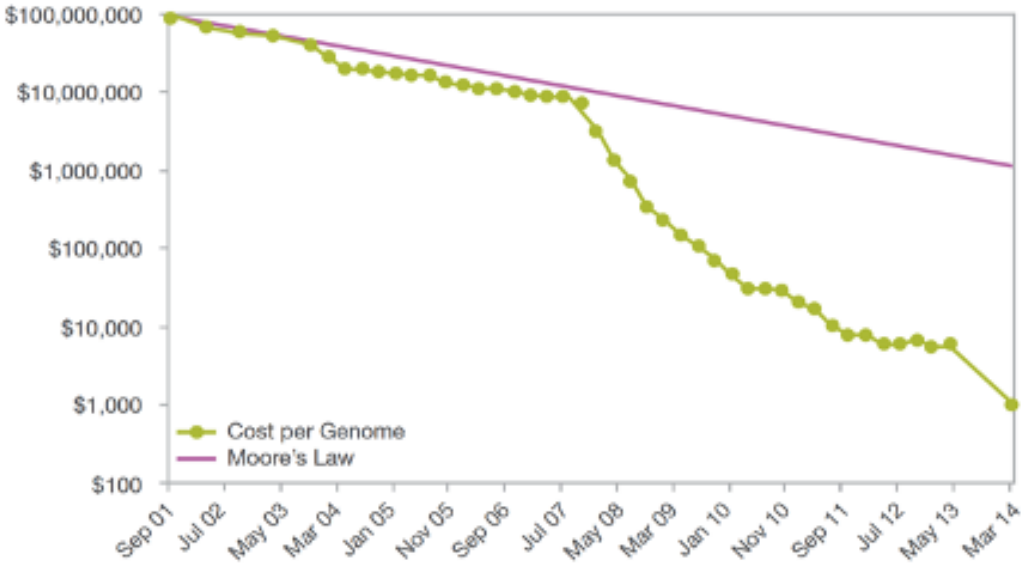
Restriction Enzymes Cut DNA at specific sites

Ligases join pieces of DNA together

Polymerases make copies of DNA (PCR)

Rate of Change in Genomics

Rate of Change



Moore's law (published in 1965) is the observation that the number of transistors in a dense integrated circuit doubles approximately every two years. (Wikipedia)

18th Century Incentives

United States patent law is authorized by the U.S. Constitution. Article One, section 8, clause 8 states:

The Congress shall have power ... To promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries;

In the 18th century, there was a need to move brains to the U.S. Legal mechanisms were designed to do this. See further link below.

Eben Moglen on Origins of Copyrights and Patents

18th Century Law in 2015

Then

- Steady and low rate of change
- Incentivise inventors by giving them exclusive rights for 20 years.
- Patents provide information for next round of inventors

Now

- Rapid increase in rate of change
- Incentivise inventors by giving them exclusive rights for 20 years.
- Information could be obsolete well before patent expires

Genomics Context

- Something like GBS was the next logical thing to try.
- The components of GBS were well known.
- Utility of GBS will last much less than 20 years.
- Therefore to maximise the utility of the method, it needs to be made freely available.

Expansion

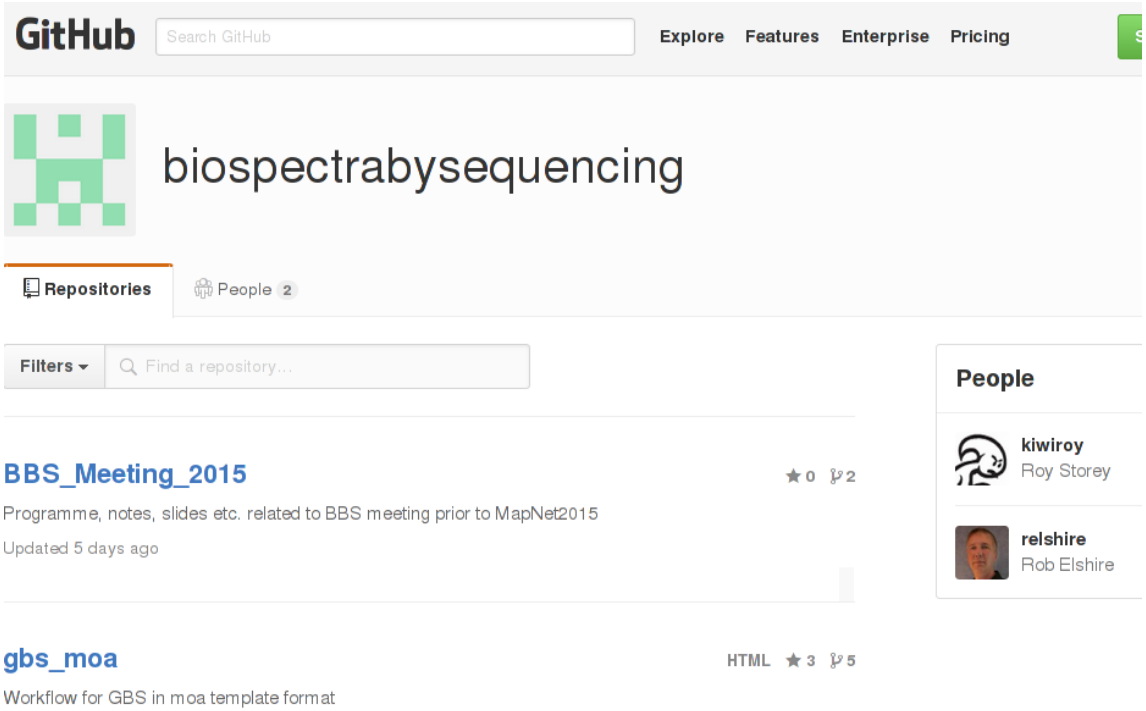
Software Tools

- Sample Tracking
- Quality control metrics
- Automation
- Reproducibility
- Downstream analyses

Information

- Species specific data
  - How many SNPs
  - Level of Multiplexing
  - Software configurations / versions
- Best practices guides
- Lab trouble shooting guides

Biospectra by Sequencing



Biospectra by Sequencing

Biospectra By Sequencing

Enabling Genomics

Log in

HomePage

The Biospectra-by-Sequencing (BBS) project is a New Zealand multi-institutional collaborative effort to build a robust genetic analysis platform initially focused on the open-source Genotyping-by-Sequencing (GBS) technology that enables more cost-effective capture of DNA-based information and to promote its adoption by a diverse range of end-users for the benefit of New Zealand's biological economy. It began as an offshoot of the MBIE and Dairy NZ funded Genomics for Production and Security in a Biological Economy (GPSBE) project at AgResearch. Currently, the project is maintained under the stewardship of The Elshire Group Limited. The initial focus of the BBS project was software development. This website is intended to complement that effort by providing a place to share practical information for use with GBS in much the same way the software is being developed and shared.

While these efforts have their origin in the New Zealand science scene, they are not NZ exclusive and we welcome like minded researchers from around the world to join us. By working together openly, we can help reduce duplication of efforts and increase our collective ability to do great research. Please join us in growing the site!

Training Videos 2014
Wet Lab
Species Specific Guidelines
Bioinformatics
FAQ?
How To Contribute

Contributors to this page: relshire .
Page last modified on Friday 11 of September, 2015 11:36:54 NZST by relshire.
The content on this page is licensed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License.

Repurposing 18th Century Law

Richard Stallman
President [Free Software Foundation](#)
[Copyleft](#)
[GNU Project](#) (1984)

Lawrence Lessig
Harvard Law Professor
[Creative Commons](#)
2001

Working Together

- The tasks require multi-disciplinary teams.
- GitHub and Wikis provide mechanisms for giving credit as well as do the work.
- FLOSS and open CC licensing are crucial.
- FLOSS and open CC licensing are not sufficient.

Not Enough, Yet

- Some BBS contributors have jobs.
- Employers hold rights to the IP on the work of employees.
- If employer removes the work from an integrated software program, then it falls over.
- Is there a way to prevent that from happening?

Contributor Agreements

debian

The Apache Software Foundation

Community-led development since 1999.

http://contributoragreements.org

Outside of the Square

- In essence, it does not require participation or workplans. But if your organisation does want to contribute, it protects all of the contributors from any one party pulling out.
- This type of agreement is *not* a negotiation between institutions, but rather defining the way your institution will interact with society.

Beyond Genomics

Closing the Gap

- Social
- Research
- Technical
- Legal

Needs and Perceptions

- MBIE's Breeding Days
  - How to accelerate plant & animal breeding
  - 'We can not move data fast enough.'
- REANNZ presentation at parliament
  - 'They built a Ferrari!'
- Central government funding a superhighway, but not investing in on-ramps

Research

- Foster trust through open, honest, and frank discussions
- Fail early and often
- International collaborations

Technical

- Harmonizing computing platforms
- Adoption of international standards
  - Compute
  - Data
  - Metadata
- Participating in international software development programmes

Legal Policy

- Currently many NZ institutions have a default IP policy that is closed. This is based on the premise that everything produced *might* have some value and therefore must not be shared.
- An alternative is to have a policy where sharing is the rule and a case must be made to not share. This is the model that has emerged in global science funding.

Science progresses one funeral at a time – I hope that it is not mine.

Questions and Discussion