

Cloud Computing and the DNA Data Race

Michael Schatz

June 8, 2011

HPDC'11/3DAPAS/ECMLS



Outline



1. Milestones in DNA Sequencing
2. Hadoop & Cloud Computing
3. Sequence Analysis in the Clouds
 1. Sequence Alignment
 2. Mapping & Genotyping
 3. Genome Assembly

Milestones in DNA Sequencing

1970

1980

1990

2000

2010

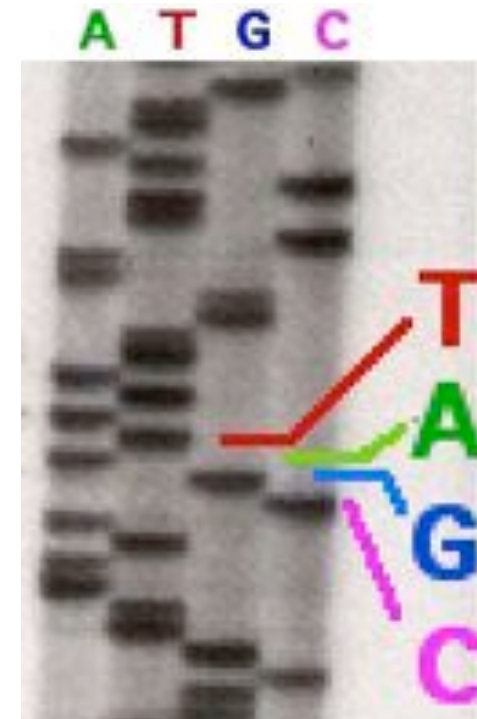
articles

Nucleotide sequence of bacteriophage ϕ X174 DNA

E. Sanger, G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, J. C. Fiddes, C. A. Hutchison III, P. M. Slesinger & M. Smith

1977

Sanger et al.
1st Complete Organism
Bacteriophage ϕ X174
5375 bp



Radioactive Chain Termination
5000bp / week / person

<http://en.wikipedia.org/wiki/File:Sequencing.jpg>
<http://www.answers.com/topic/automated-sequencer>

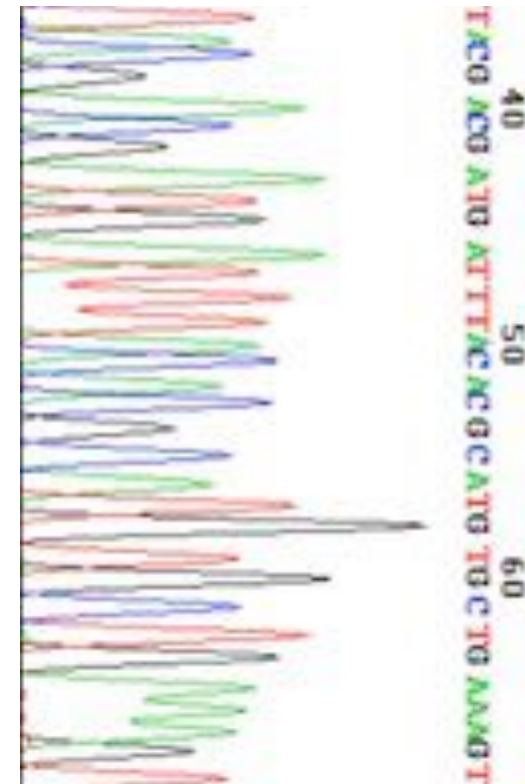
Milestones in DNA Sequencing



1987

Applied Biosystems markets the ABI 370 as the first automated sequencing machine

http://commons.wikimedia.org/wiki/File:370A_automated_DNA_sequencer.jpg



Fluorescent Dye Termination
350bp / lane x 16 lanes =
5600bp / day / machine

<http://www.answers.com/topic/automated-sequencer>

Milestones in DNA Sequencing



1995

Fleischmann *et al.*
1st Free Living Organism
TIGR Assembler. 1.8Mbp



2000

Myers *et al.*
1st Large WGS Assembly.
Celera Assembler. 116 Mbp



2001

Venter *et al.*,
Human Genome
Celera Assembler. 2.9 Gbp

ABI 3700: 500 bp reads x 768 samples / day = 384,000 bp / day.

"The machine was so revolutionary that it could decode in a single day the same amount of genetic material that most DNA labs could produce in a year." J. Craig Venter

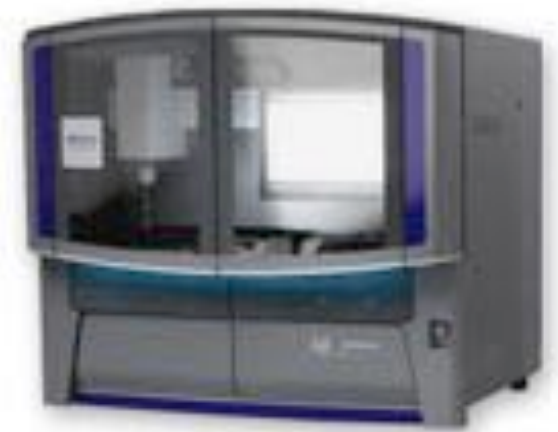
Milestones in DNA Sequencing



2004
454/Roche
Pyrosequencing
Current Specs (Titanium):
1M 400bp reads / run =
1Gbp / day

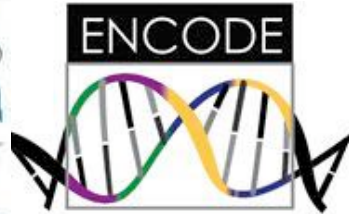


2007
Illumina
Sequencing by Synthesis
Current Specs (HiSeq 2000):
2.5B 100bp reads / run =
60Gbp / day

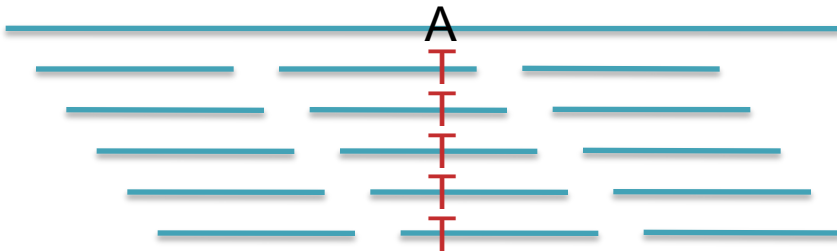


2008
ABI / Life Technologies
SOLiD Sequencing
Current Specs (5500xl):
5B 75bp reads / run =
30Gbp / day

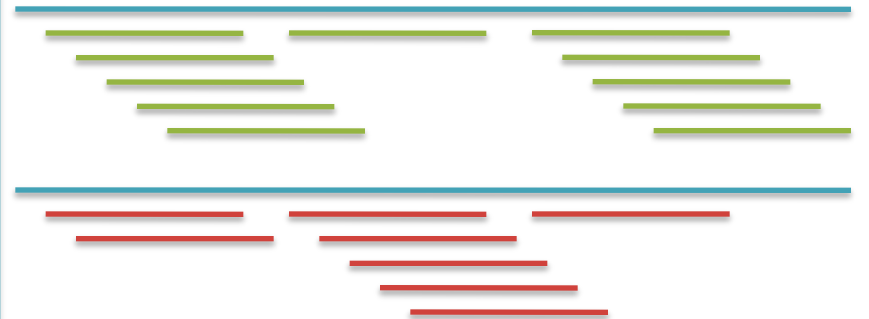
Second Generation Sequencing Applications



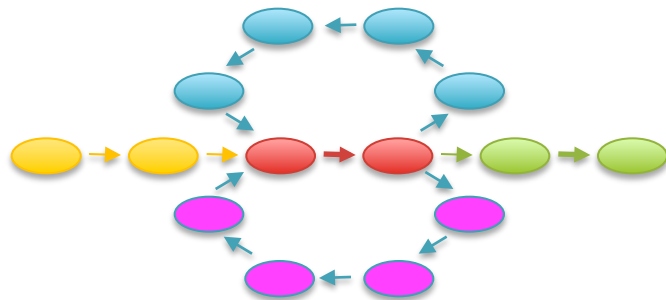
Alignment & Variations



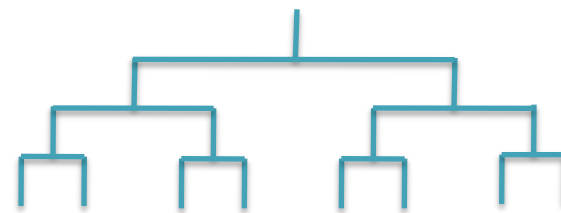
Differential Analysis



De novo Assembly



Phylogeny & Evolution



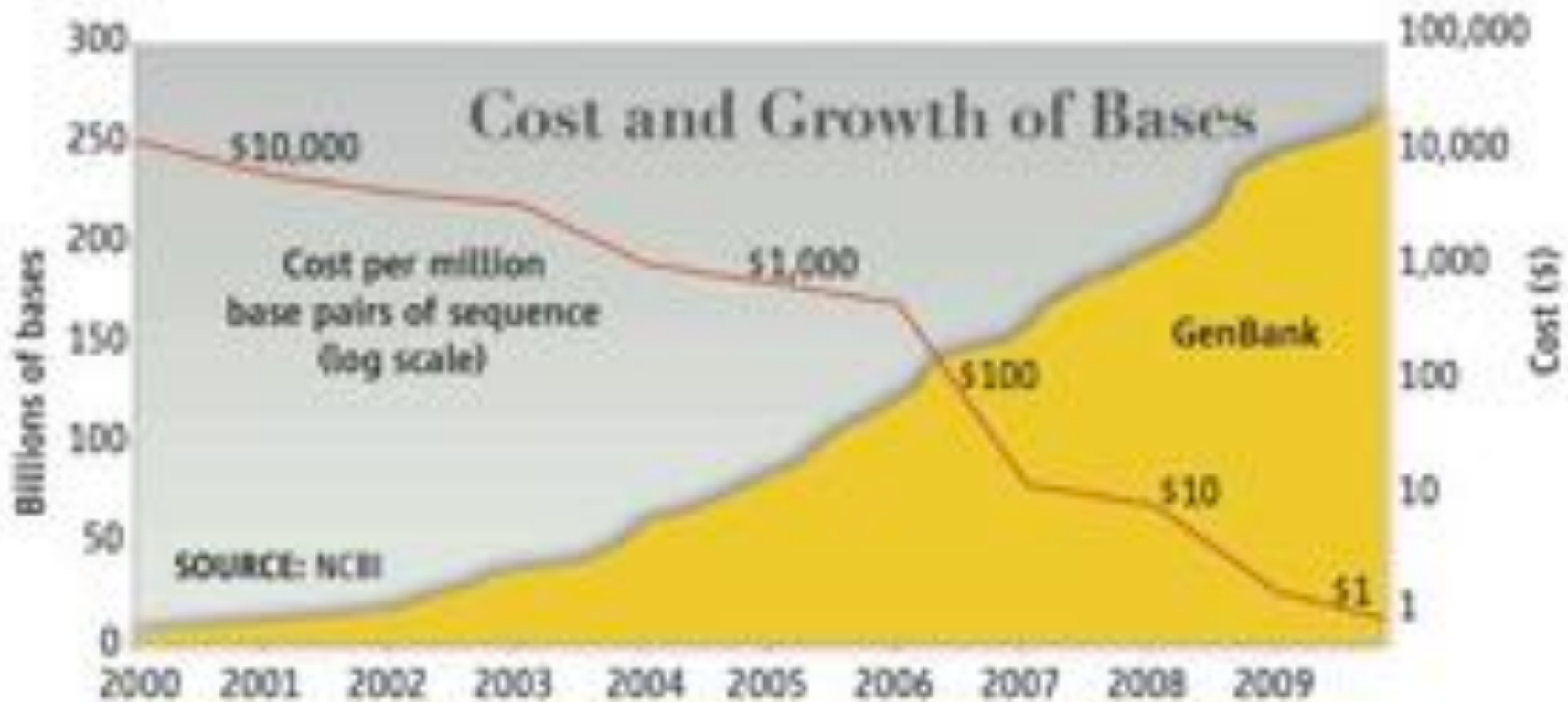
Sequencing Centers



Next Generation Genomics: World Map of High-throughput Sequencers
<http://pathogenomics.bham.ac.uk/hts/>

The DNA Data Tsunami

Current world-wide sequencing capacity exceeds 33Tbp/day (12Pbp/year) and is growing at 5x per year!



"Will Computers Crash Genomics?"

Elizabeth Pennisi (2011) *Science*. 331(6018): 666-668.

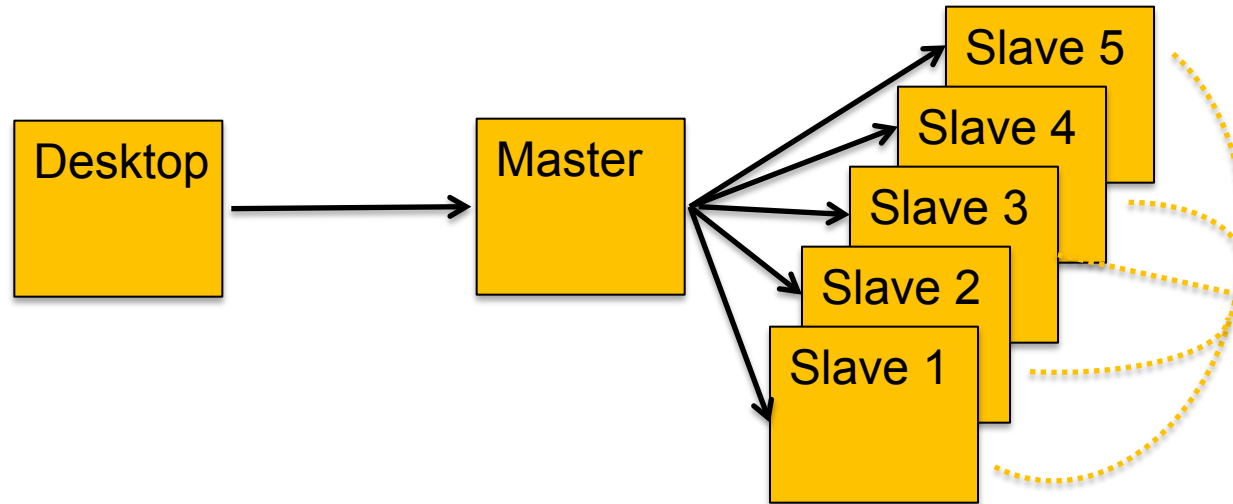
Hadoop MapReduce

<http://hadoop.apache.org>

- MapReduce is Google's framework for large data computations
 - Data and computations are spread over thousands of computers
 - Indexing the Internet, PageRank, Machine Learning, etc... (Dean and Ghemawat, 2004)
 - 946 PB processed in May 2010 (Jeff Dean at Stanford, 11.10.2010)
 - Hadoop is the leading open source implementation
 - Developed and used by Yahoo, Facebook, Twitter, Amazon, etc
 - GATK is an alternative implementation specifically for NGS
- Benefits
 - Scalable, Efficient, Reliable
 - Easy to Program
 - Runs on commodity computers
- Challenges
 - Redesigning / Retooling applications
 - Not Condor, Not MPI
 - Everything in MapReduce



System Architecture



- Hadoop Distributed File System (HDFS)
 - Data files partitioned into large chunks (64MB), replicated on multiple nodes
 - Computation moves to the data, rack-aware scheduling
- Hadoop MapReduce system won the 2009 GreySort Challenge
 - Sorted 100 TB in 173 min (578 GB/min) using 3452 nodes and 4x3452 disks
 - Provides many disks in addition to many cores

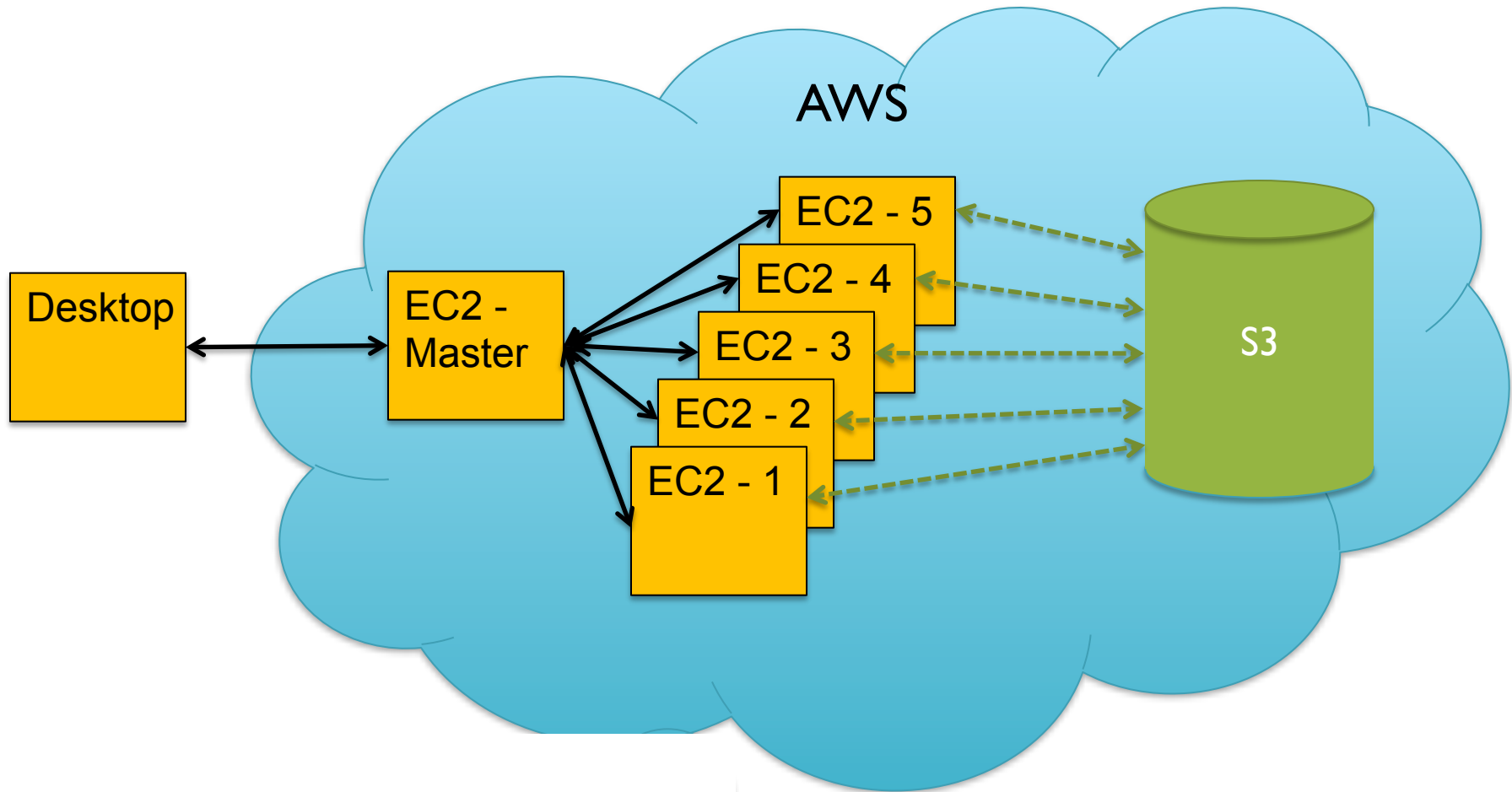
Amazon Web Services

<http://aws.amazon.com>

- All you need is a credit card, and you can immediately start using one of the largest datacenters in the world
- Elastic Compute Cloud (EC2)
 - Flexible allocation of virtual machines
 - Pricing starting at 2¢ / hour
- Simple Storage Service (S3)
 - Pricing starts at 15¢ / GB / month
 - 5.5¢ / GB / month for over 5 PB
- Plus many others



Hadoop on AWS

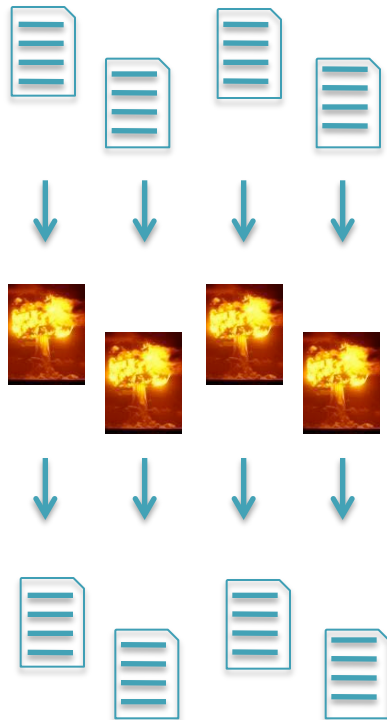


If you don't have 1000s of machines, rent them from Amazon

- After machines pool up, ssh to master as if it was a local machine.
- Use S3 for persistent data storage, with very fast interconnect to EC2.

Programming Models

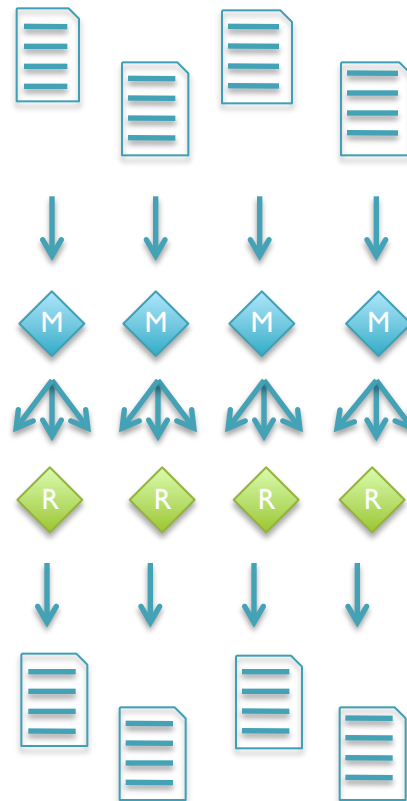
Embarrassingly Parallel



Map-only

Each item is Independent
Traditional Batch Computing

Loosely Coupled



MapReduce

Independent-Shuffle-Independent
Batch Computing + Data Exchange

Tightly Coupled

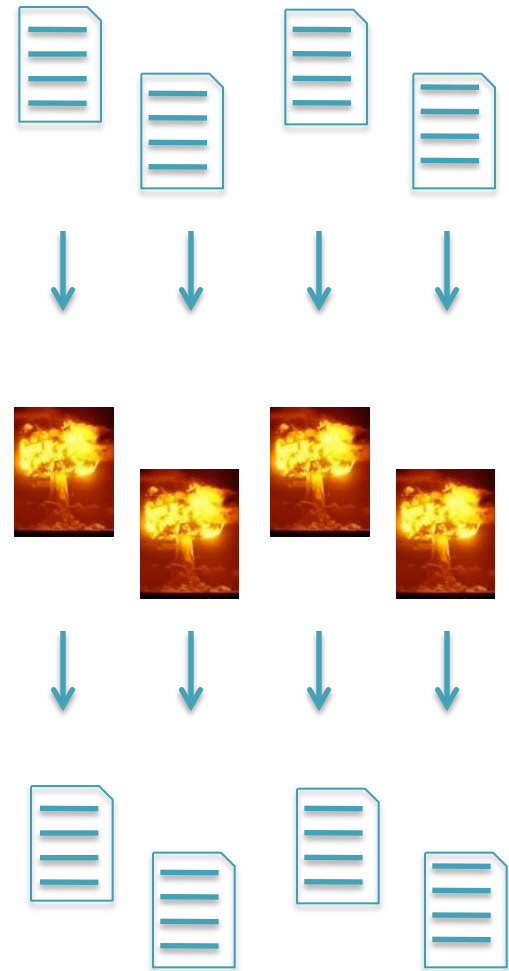


Iterative MapReduce

Nodes interact with other nodes
Big Data MPI

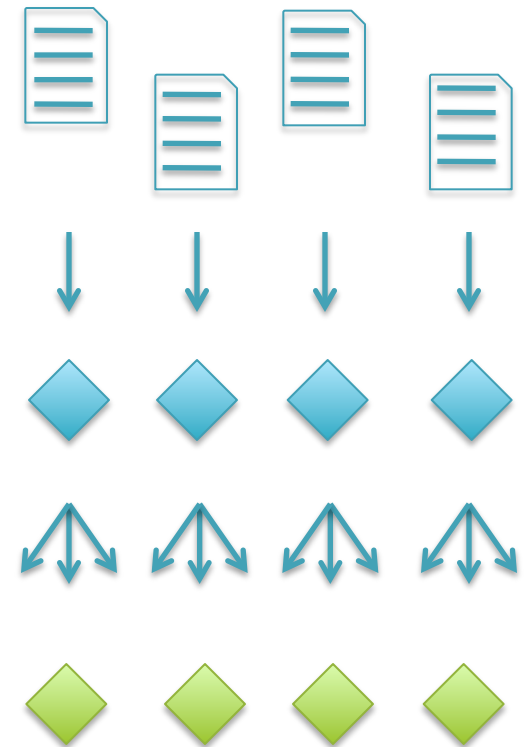
I. Embarrassingly Parallel

- Batch computing
 - Each item is independent
 - Split input into many chunks
 - Process each chunk separately on a different computer
- Challenges
 - Distributing work, load balancing, monitoring & restart
- Technologies
 - Condor, Sun Grid Engine
 - Amazon Simple Queue

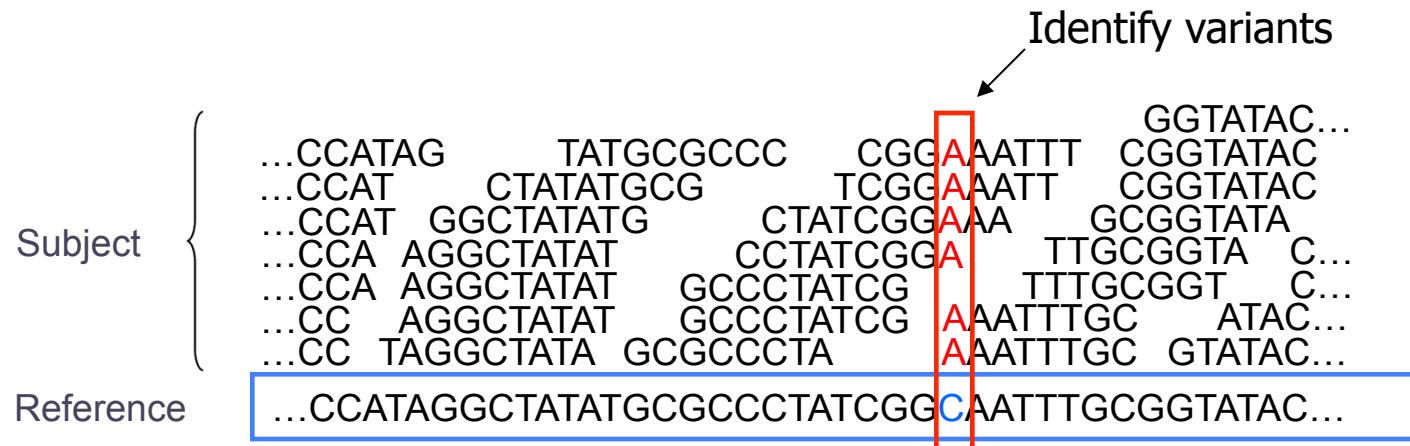


2. Loosely Coupled

- Divide and conquer
 - Independently process many items
 - Group partial results
 - Scan partial results into final answer
- Challenges
 - Batch computing challenges
 - + Shuffling of huge datasets
- Technologies
 - Hadoop, Elastic MapReduce, Dryad
 - Parallel Databases



Short Read Mapping



- Given a reference and many subject reads, report one or more “good” end-to-end alignments per alignable read
 - Find where the read most likely originated
 - Fundamental computation for many assays
 - Genotyping RNA-Seq Methyl-Seq
 - Structural Variations Chip-Seq Hi-C-Seq

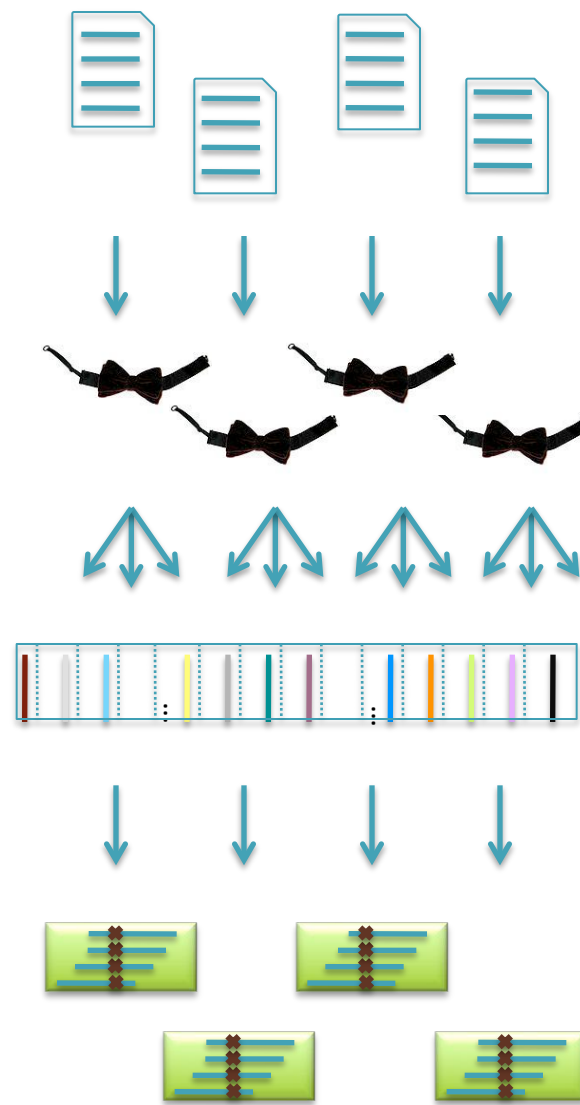
- Desperate need for scalable solutions
 - Single human requires >1,000 CPU hours / genome



Crossbow

<http://bowtie-bio.sourceforge.net/crossbow>

- Align billions of reads and find SNPs
 - Reuse software components: Hadoop Streaming
- Map: Bowtie (Langmead *et al.*, 2009)
 - Find best alignment for each read
 - Emit (chromosome region, alignment)
- Shuffle: Hadoop
 - Group and sort alignments by region
- Reduce: SOAPsnp (Li *et al.*, 2009)
 - Scan alignments for divergent columns
 - Accounts for sequencing error, known SNPs



Performance in Amazon EC2

<http://bowtie-bio.sourceforge.net/crossbow>

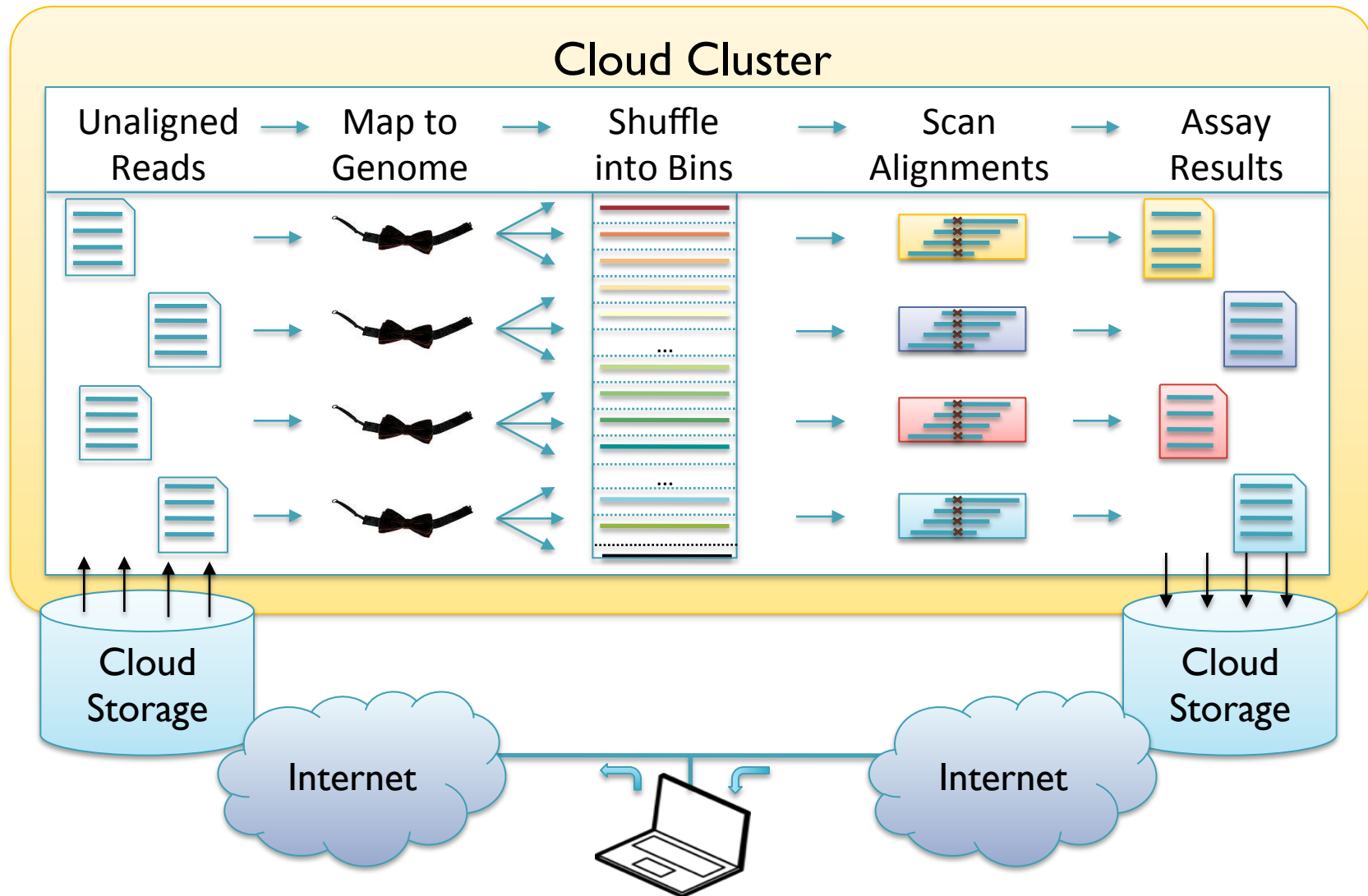
	Asian Individual Genome		
Data Loading	3.3 B reads	106.5 GB	\$10.65
Data Transfer	1h :15m	40 cores	\$3.40
Setup	0h : 15m	320 cores	\$13.94
Alignment	1h : 30m	320 cores	\$41.82
Variant Calling	1h : 00m	320 cores	\$27.88
End-to-end	4h : 00m		\$97.69

Discovered 3.7M SNPs in one human genome for ~\$100 in an afternoon.
Accuracy validated at >99%

Searching for SNPs with Cloud Computing.

Langmead B, Schatz MC, Lin J, Pop M, Salzberg SL (2009) *Genome Biology*. **10**:R134

Map-Shuffle-Scan for Genomics

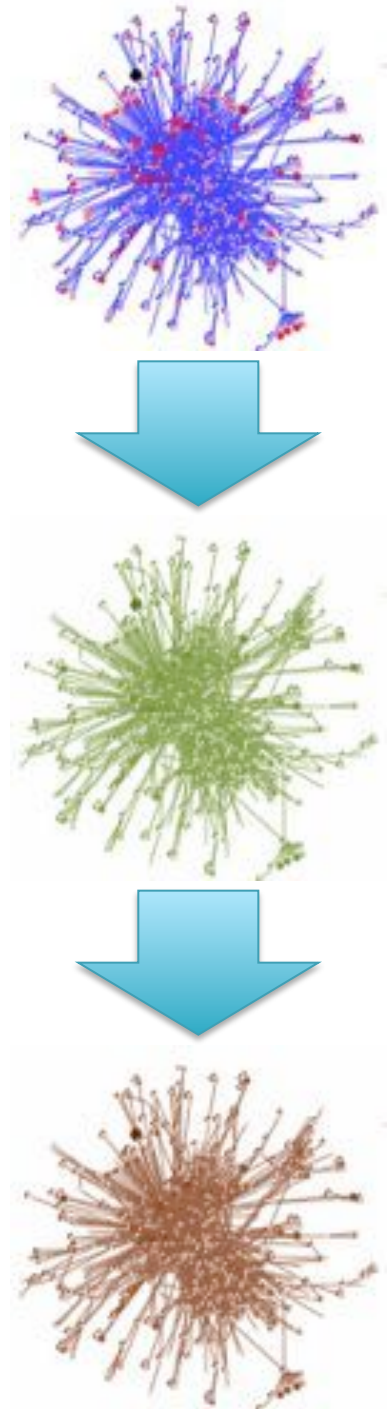


Cloud Computing and the DNA Data Race.

Schatz, MC, Langmead B, Salzberg SL (2010) *Nature Biotechnology*. **28**:691-693

3. Tightly Coupled

- Computation that cannot be partitioned
 - Graph Analysis
 - Molecular Dynamics
 - Population simulations
- Challenges
 - Loosely coupled challenges
 - + Parallel algorithms design
- Technologies
 - MPI
 - MapReduce, Dryad, Pregel

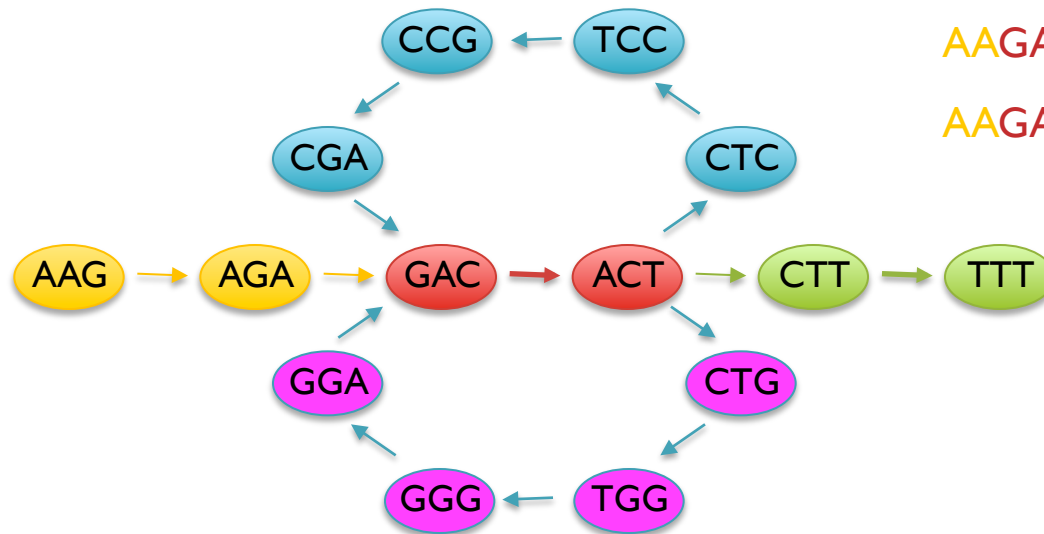


Short Read Assembly

Reads

AAGA
ACTT
ACTC
ACTG
AGAG
CCGA
CGAC
CTCC
CTGG
CTTT
...

de Bruijn Graph



Potential Genomes

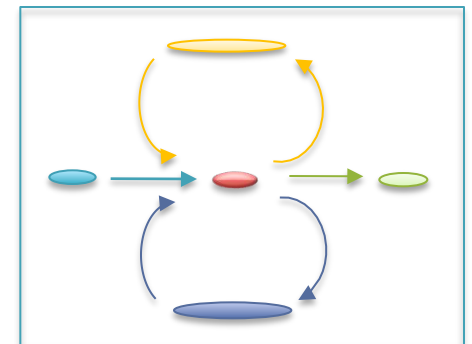
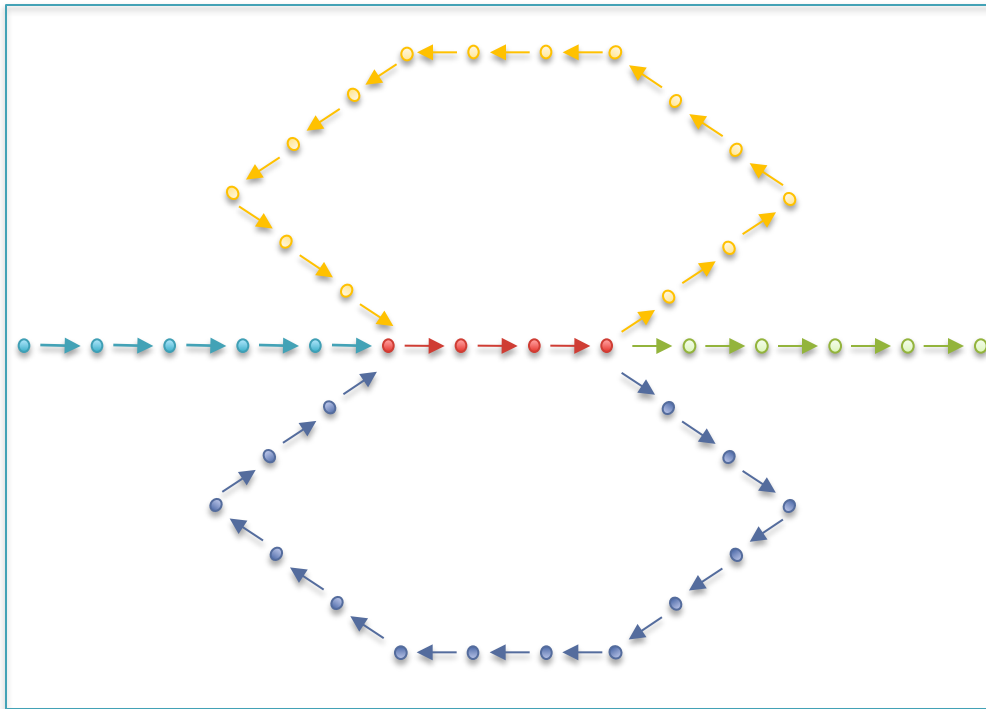
AAGACTCCGACTGGGACTTT

AAGACTGGGACTCCGACTTT

- Genome assembly as finding an Eulerian tour of the de Bruijn graph
 - Human genome: >3B nodes, >10B edges
- The new short read assemblers require tremendous computation
 - Velvet (Zerbino & Birney, 2008) serial: > 2TB of RAM
 - ABySS (Simpson *et al.*, 2009) MPI: 168 cores x ~96 hours
 - SOAPdenovo (Li *et al.*, 2010) pthreads: 40 cores x 40 hours, >140 GB RAM

Graph Compression

- After construction, many edges are unambiguous
 - Merge together compressible nodes
 - Graph physically distributed over hundreds of computers



Warmup Exercise

- Who here was born closest to June 8?
 - You can only compare to 1 other person at a time



Find winner among 64 teams in just 6 rounds

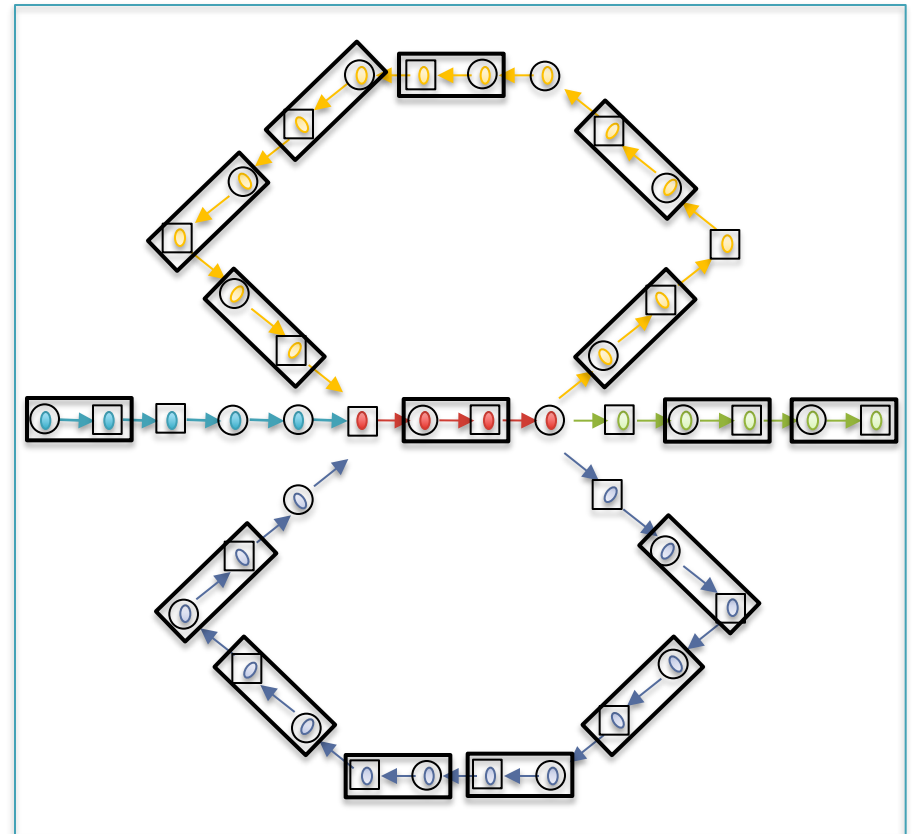
Fast Path Compression

Challenges

- Nodes stored on different computers
- Nodes can only access direct neighbors

Randomized List Ranking

- Randomly assign \textcircled{H} / $\square T$ to each compressible node
- Compress $\textcircled{H} \rightarrow \square T$ links



Initial Graph: 42 nodes

Randomized Speed-ups in Parallel Computation.

Vishkin U. (1984) *ACM Symposium on Theory of Computation*. 230-239.

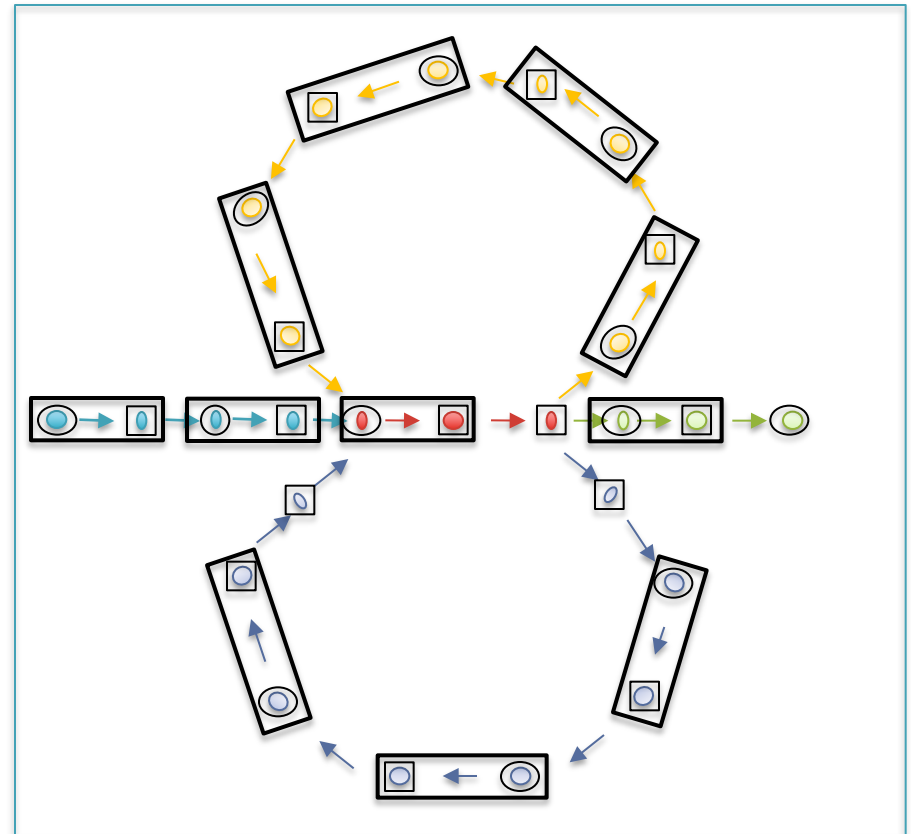
Fast Path Compression

Challenges

- Nodes stored on different computers
- Nodes can only access direct neighbors

Randomized List Ranking

- Randomly assign \textcircled{H} / $\square T$ to each compressible node
- Compress $\textcircled{H} \rightarrow \square T$ links



Round 1: 26 nodes (38% savings)

Randomized Speed-ups in Parallel Computation.

Vishkin U. (1984) *ACM Symposium on Theory of Computation*. 230-239.

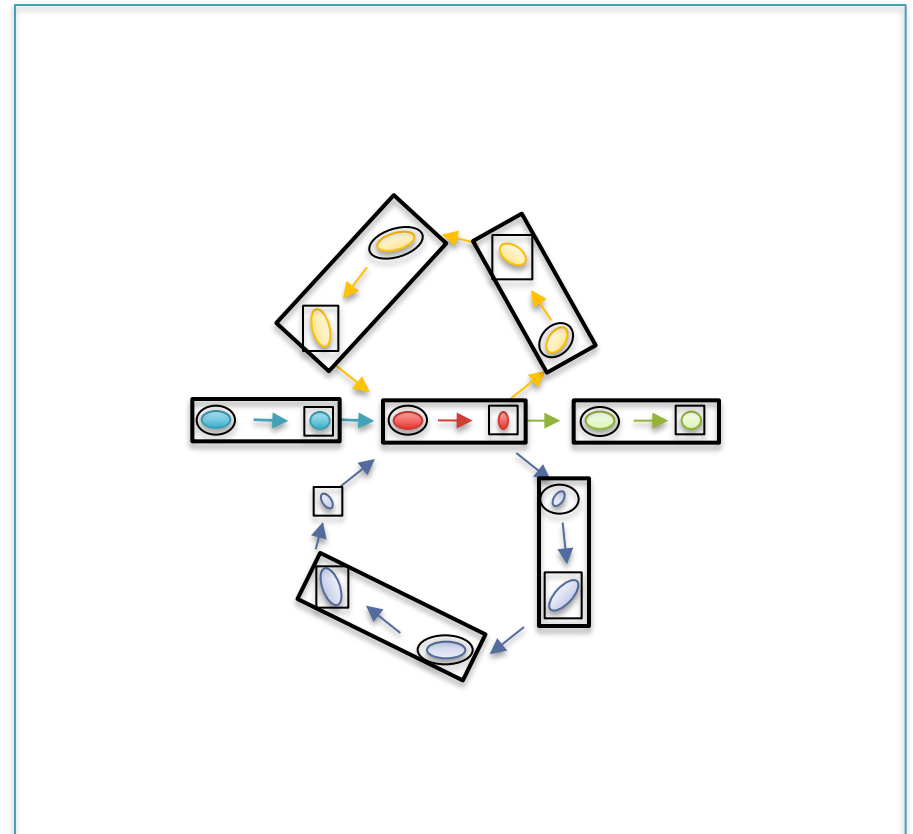
Fast Path Compression

Challenges

- Nodes stored on different computers
- Nodes can only access direct neighbors

Randomized List Ranking

- Randomly assign $\textcircled{\text{H}}$ / $\boxed{\text{T}}$ to each compressible node
- Compress $\textcircled{\text{H}} \rightarrow \boxed{\text{T}}$ links



Round 2: 15 nodes (64% savings)

Randomized Speed-ups in Parallel Computation.

Vishkin U. (1984) *ACM Symposium on Theory of Computation*. 230-239.

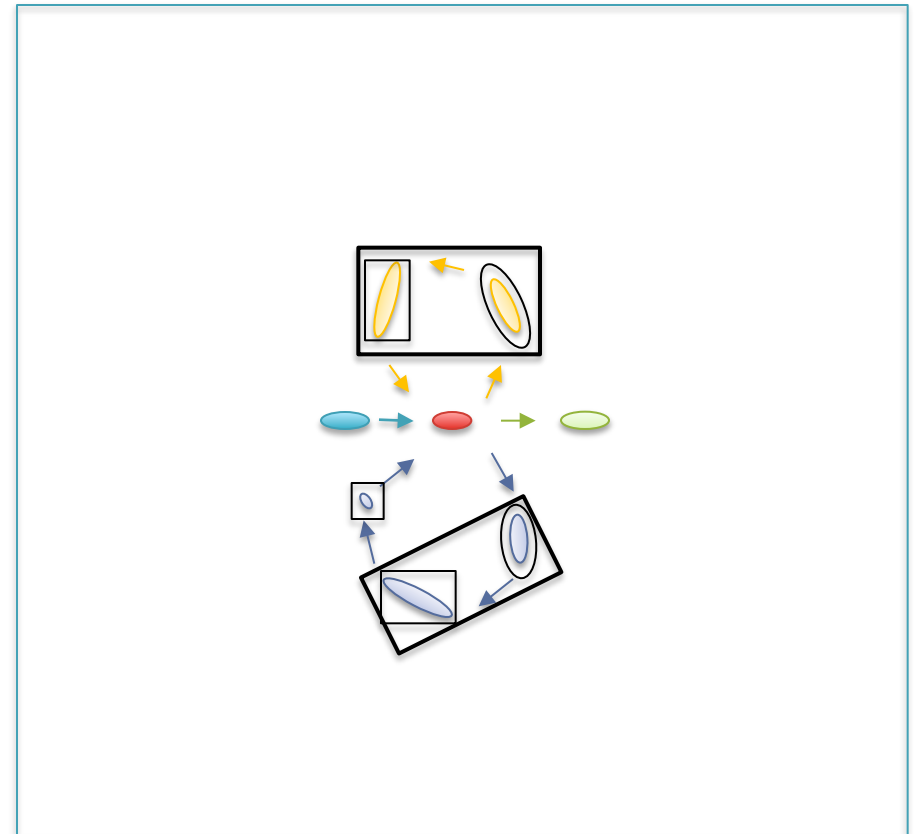
Fast Path Compression

Challenges

- Nodes stored on different computers
- Nodes can only access direct neighbors

Randomized List Ranking

- Randomly assign $\textcircled{\text{H}}$ / $\boxed{\text{T}}$ to each compressible node
- Compress $\textcircled{\text{H}} \rightarrow \boxed{\text{T}}$ links



Round 2: 8 nodes (81% savings)

Randomized Speed-ups in Parallel Computation.

Vishkin U. (1984) *ACM Symposium on Theory of Computation*. 230-239.

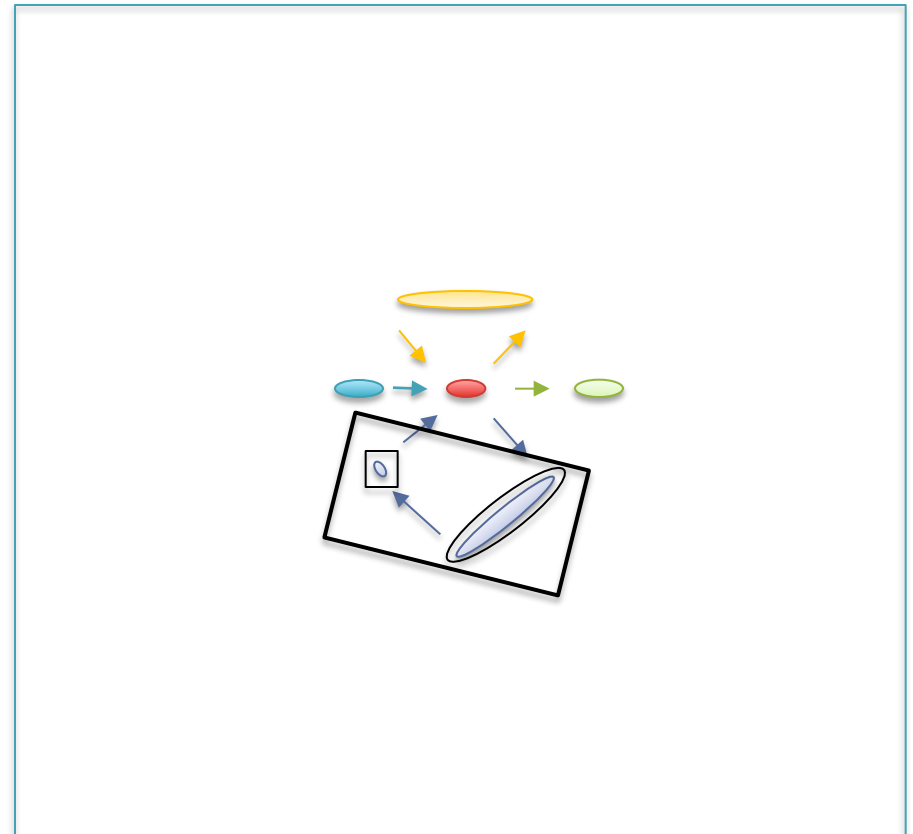
Fast Path Compression

Challenges

- Nodes stored on different computers
- Nodes can only access direct neighbors

Randomized List Ranking

- Randomly assign $\textcircled{\text{H}}$ / $\boxed{\text{T}}$ to each compressible node
- Compress $\textcircled{\text{H}} \rightarrow \boxed{\text{T}}$ links



Round 3: 6 nodes (86% savings)

Randomized Speed-ups in Parallel Computation.

Vishkin U. (1984) *ACM Symposium on Theory of Computation*. 230-239.

Fast Path Compression

Challenges

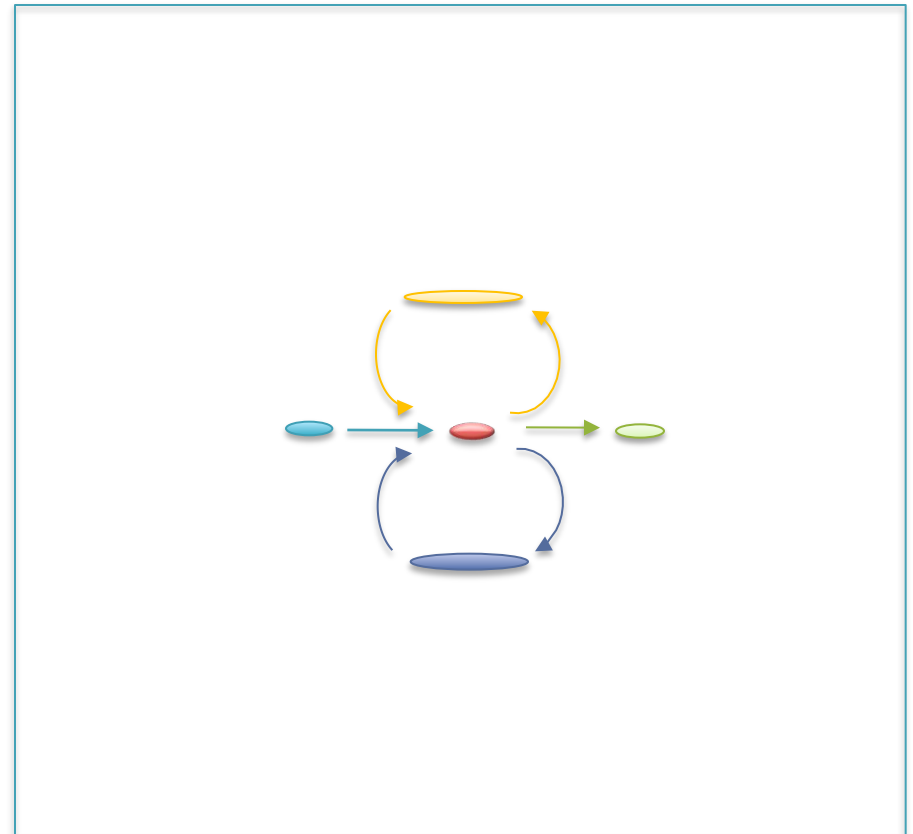
- Nodes stored on different computers
- Nodes can only access direct neighbors

Randomized List Ranking

- Randomly assign $\textcircled{\text{H}}$ / $\boxed{\text{T}}$ to each compressible node
- Compress $\textcircled{\text{H}} \rightarrow \boxed{\text{T}}$ links

Performance

- Compress all chains in $\log(S)$ rounds



Round 4: 5 nodes (88% savings)

Randomized Speed-ups in Parallel Computation.

Vishkin U. (1984) *ACM Symposium on Theory of Computation*. 230-239.

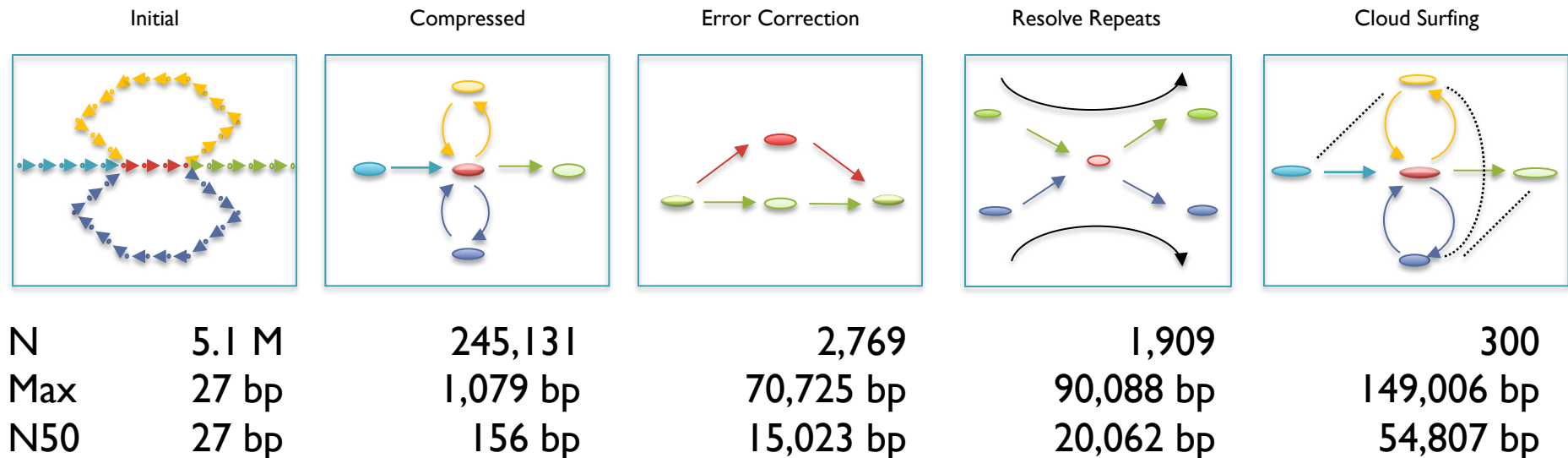
Contrail

<http://contrail-bio.sourceforge.net>



De novo bacterial assembly

- *Genome: E. coli* K12 MGI655, 4.6Mbp
- *Input: 20.8M* 36bp reads, 200bp insert (~150x coverage)
- *Preprocessor: Quake Error Correction*



Assembly of Large Genomes with Cloud Computing.

Schatz MC, Sommer D, Kelley D, Pop M, et al. *In Preparation.*

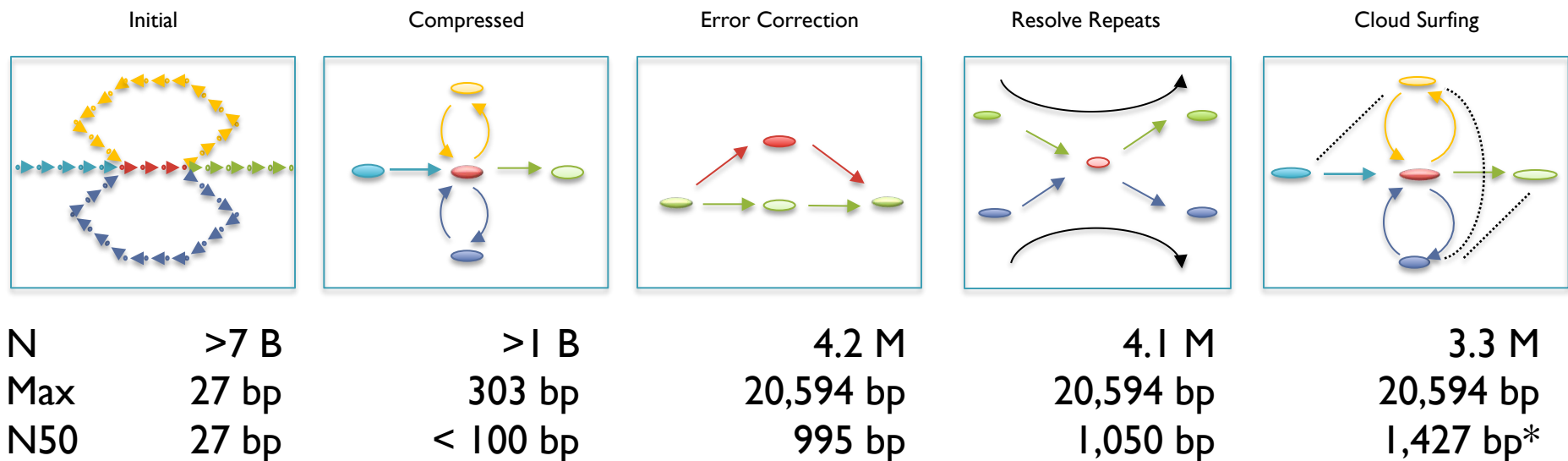
Contrail

<http://contrail-bio.sourceforge.net>



De novo Assembly of the Human Genome

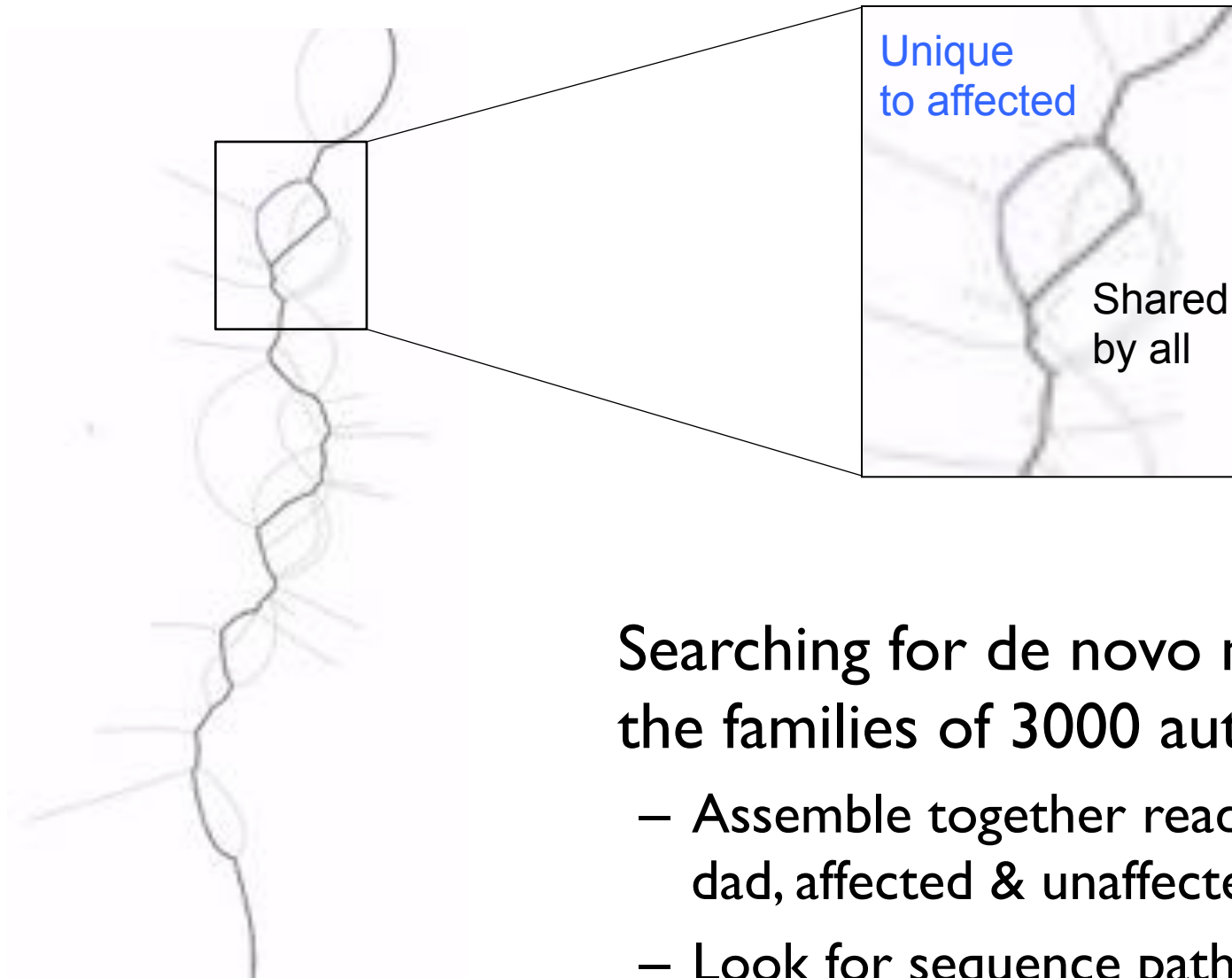
- *Genome*: African male NAI8507 (SRA000271, Bentley *et al.*, 2008)
- *Input*: 3.5B 36bp reads, 210bp insert (~40x coverage)



Assembly of Large Genomes with Cloud Computing.

Schatz MC, Sommer D, Kelley D, Pop M, *et al.* *In Preparation.*

De novo mutations and de Bruijn Graphs



COLEC12
C->A

Searching for de novo mutations in the families of 3000 autistic children.

- Assemble together reads from mom, dad, affected & unaffected children
- Look for sequence paths unique to affected child

Hadoop for NGS Analysis



CloudBurst

Highly Sensitive Short Read Mapping with MapReduce

100x speedup mapping on 96 cores @ Amazon

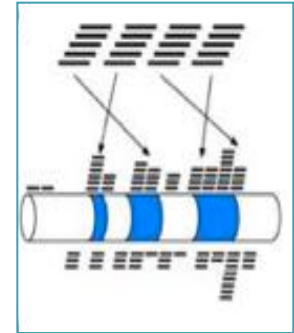
<http://cloudburst-bio.sf.net>

(Schatz, 2009)

Myrna

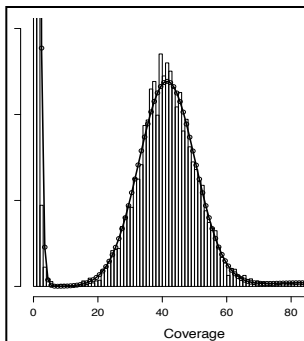
Cloud-scale differential gene expression for RNA-seq

Expression of 1.1 billion RNA-Seq reads in ~2 hours for ~\$66



(Langmead, Hansen, Leek, 2010)

<http://bowtie-bio.sf.net/myrna/>



Quake

Quality-aware error correction of short reads

Correct 97.9% of errors with 99.9% accuracy

<http://www.cbcb.umd.edu/software/quake/>

(Kelley, Schatz, Salzberg, 2010)

Genome Indexing

Rapid Parallel Construction of Genome Index

Construct the BWT of the human genome in 9 minutes

```
$GATTACA  
A$GATTAC  
ACA$GATT  
ATTACA$G  
CA$GATTA  
GATTACA£  
TACA$GAT  
TTACA$GA
```

(Menon, Bhat, Schatz, 2011*)

<http://code.google.com/p/genome-indexing/>



Summary

- Staying afloat in the data deluge means computing in parallel
 - Hadoop + Cloud computing is an attractive platform for large scale sequence analysis and computation
- Significant obstacles ahead
 - Price
 - Transfer time
 - Privacy / security requirements
 - Time and expertise required for development
- Emerging technologies are a great start, but we need continued research
 - Need integration across disciplines
 - A word of caution: new technologies are new

Acknowledgements

CSHL

Dick McCombie

Melissa delaBastide

Mike Wigler

Ivan Iossifov

Zach Lippman

Doreen Ware

Mitch Bekritsky

SBU

Steve Skiena

Matt Titmus

Rohith Menon

Goutham Bhat

Hayan Lee

JHU

Steven Salzberg

Ben Langmead

Jeff Leek

Univ. of Maryland

Mihai Pop

Art Delcher

Jimmy Lin

Adam Phillippy

David Kelley

Dan Sommer



Beyond the Genome 2011

19 - 22 September 2011, Washington DC

In association with



Genome **Biology**



Genome **Medicine**

Thank You!

<http://schatzlab.cshl.edu>

@mike_schatz