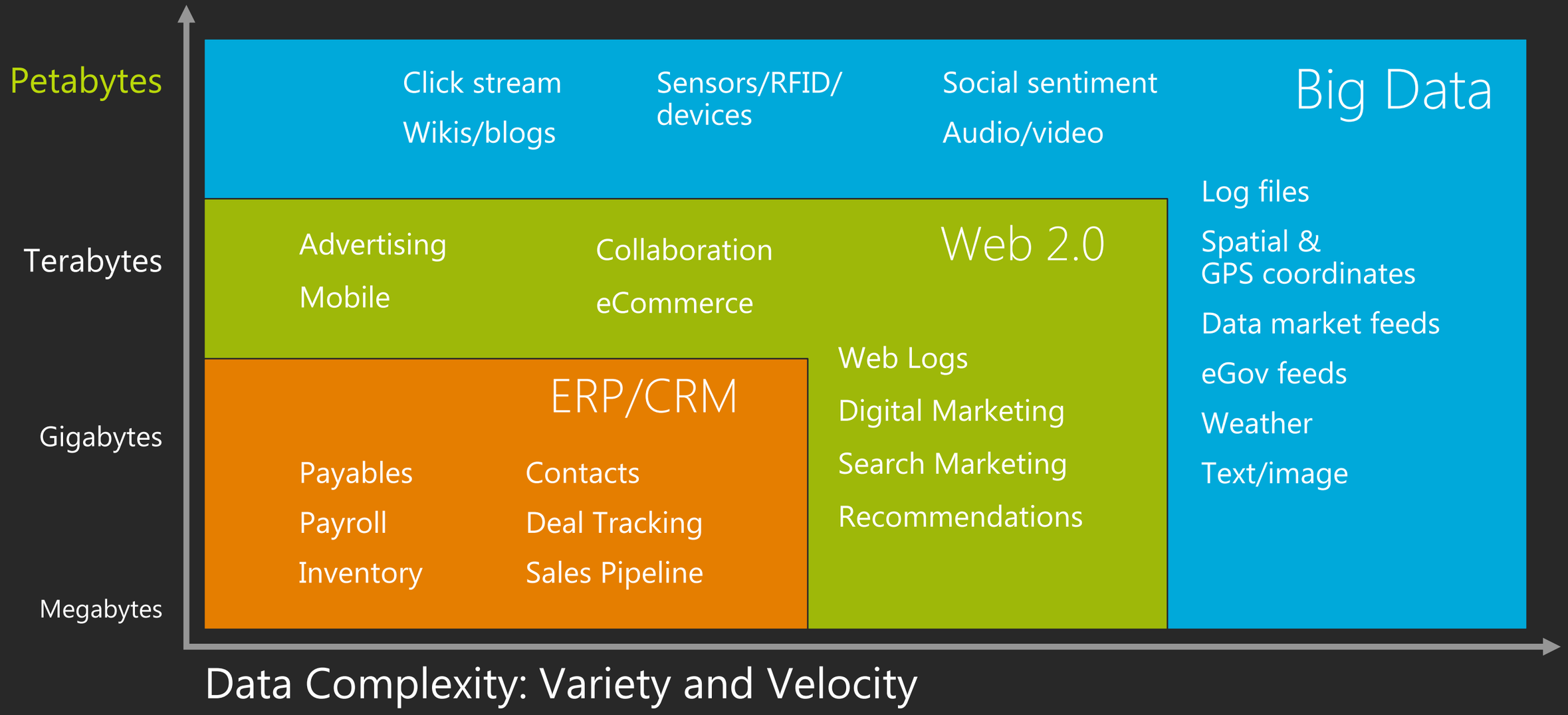
ANALYZING BIG DATA

WHAT IS BIG DATA?



WHAT IS BIG DATA AND WHY NOW?

Volume
Variety
Velocity

Data explosion

Changing economics

Hadoop

Cheap, Distributed Storage & Processing

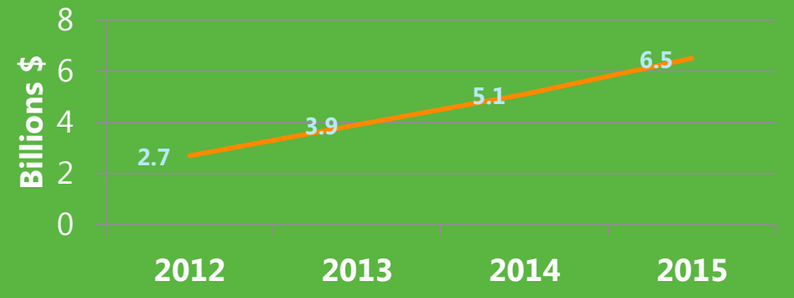
> By 2015, organizations that build a modern information management system will outperform their peers financially by 20 percent.
>
> – Gartner, Mark Beyer, "Information Management in the 21st Century"

BIG DATA IS A GROWTH OPPORTUNITY FOR CLIENTS

Big Data is a Big Priority for Clients

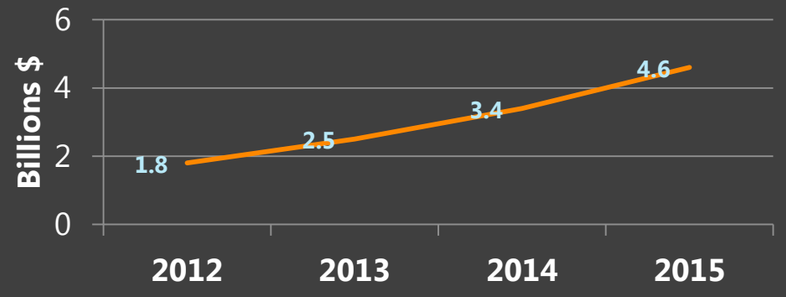
49% of top CEOs and CIOs are currently using Big Data for customer analytics[1]

Big Data Services Growth



39% compound annual growth rate2

Big Data Software Growth

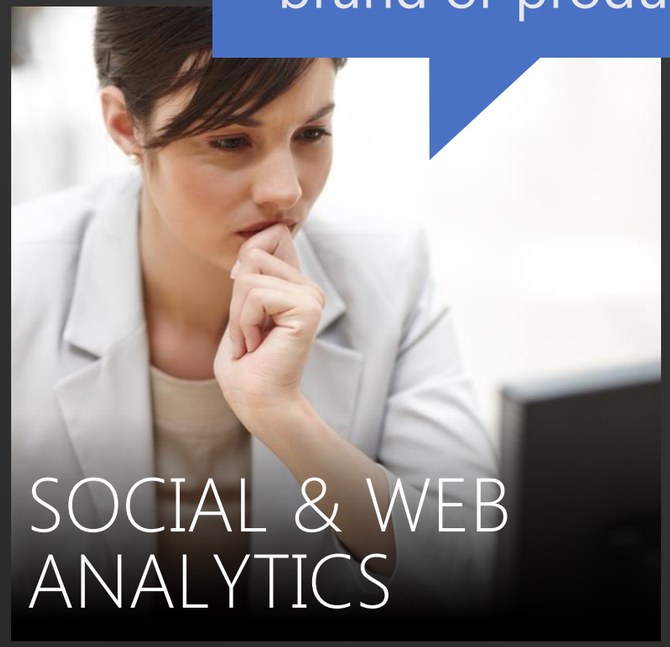


34% compound annual growth rate[2]

1. McKinsey&Company, McKinsey Global Survey Results, Minding Your Digital Business, 2012
2. IDC Market Analysis, Worldwide Big Data Technology and Services 2012–2015 Forecast , 2012

A NEW SET OF QUESTIONS

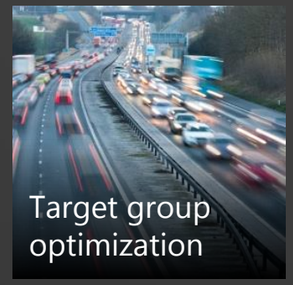
What's the market sentiment for my brand or products

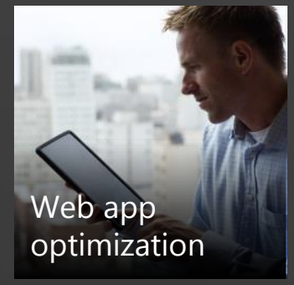


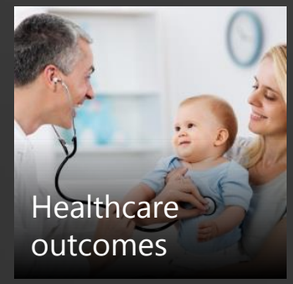
How do I optimize my target group based on demographics and health patterns?

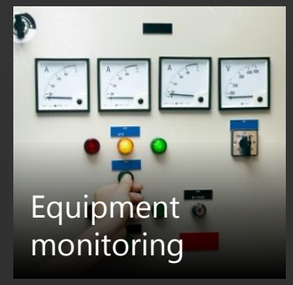
How do I better predict future outcomes?

NEW OPPORTUNITIES









BIG DATA REQUIRES AN END-TO-END APPROACH

DATA MANAGEMENT

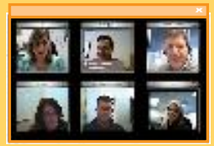
- Relational
- Non-relational
- Analytical
- Streaming

DATA ENRICHMENT

- Discover
- Combine
- Refine

INSIGHT

Self-Service

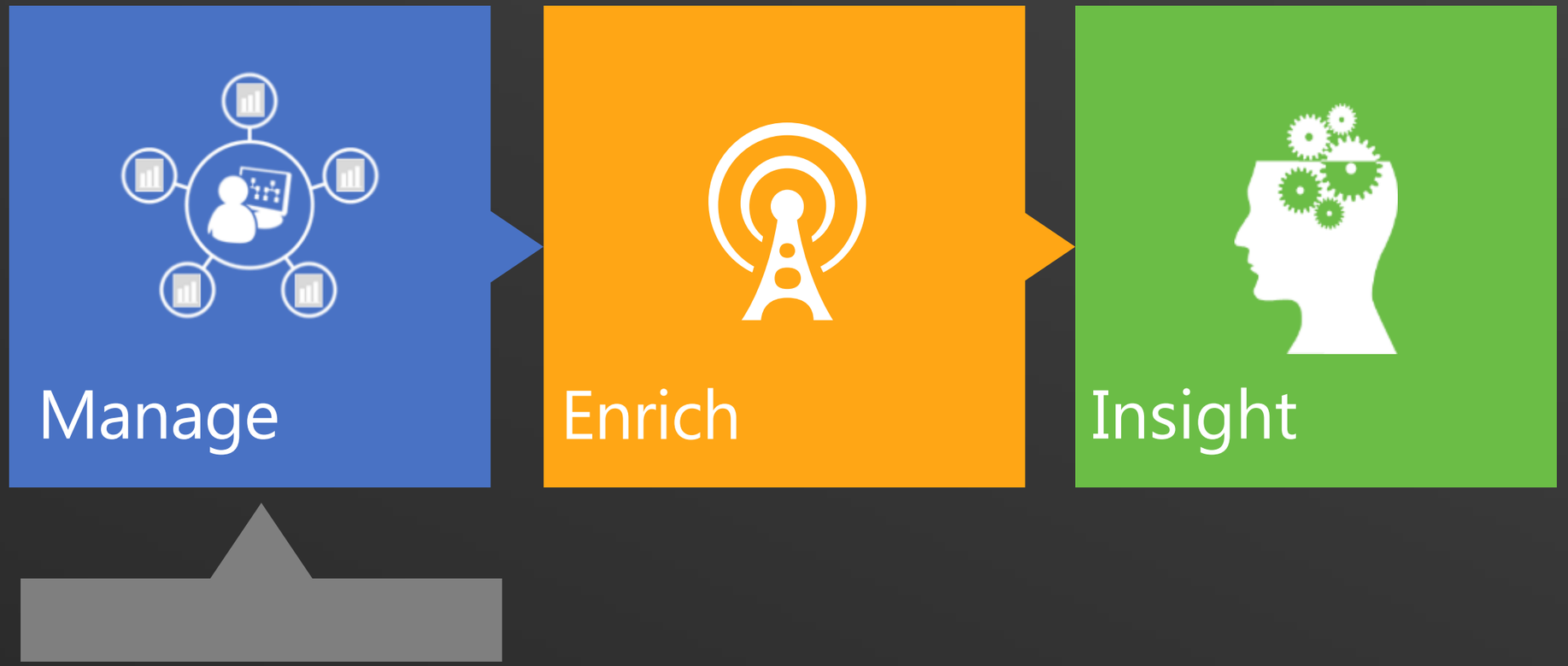


Collaboration

Corporate Apps

Devices

THE BIG DATA LIFECYCLE



MANAGE ANY DATA, ANY SIZE, ANYWHERE

e.g. STRUCTURED & UNSTRUCTURED DATA

Unified Monitoring, Management & Security

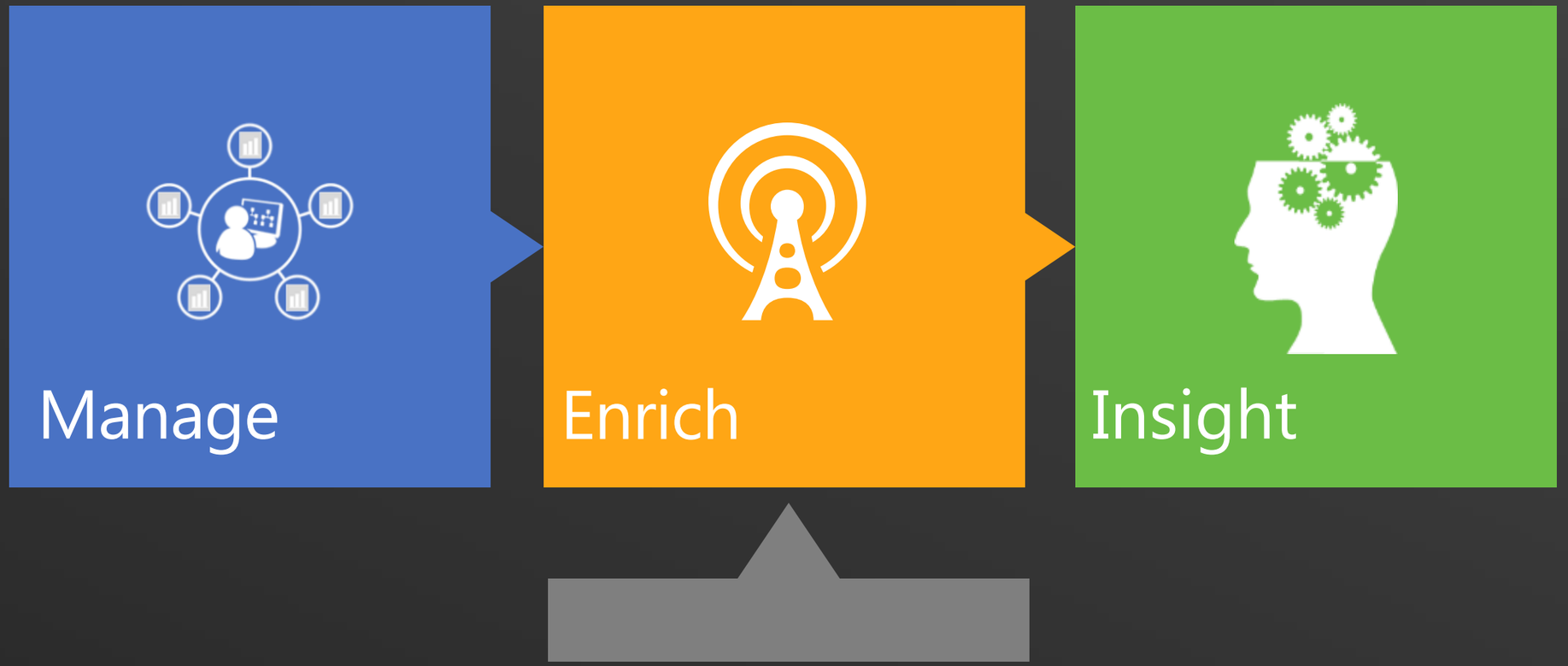
Relational

Non-Relational

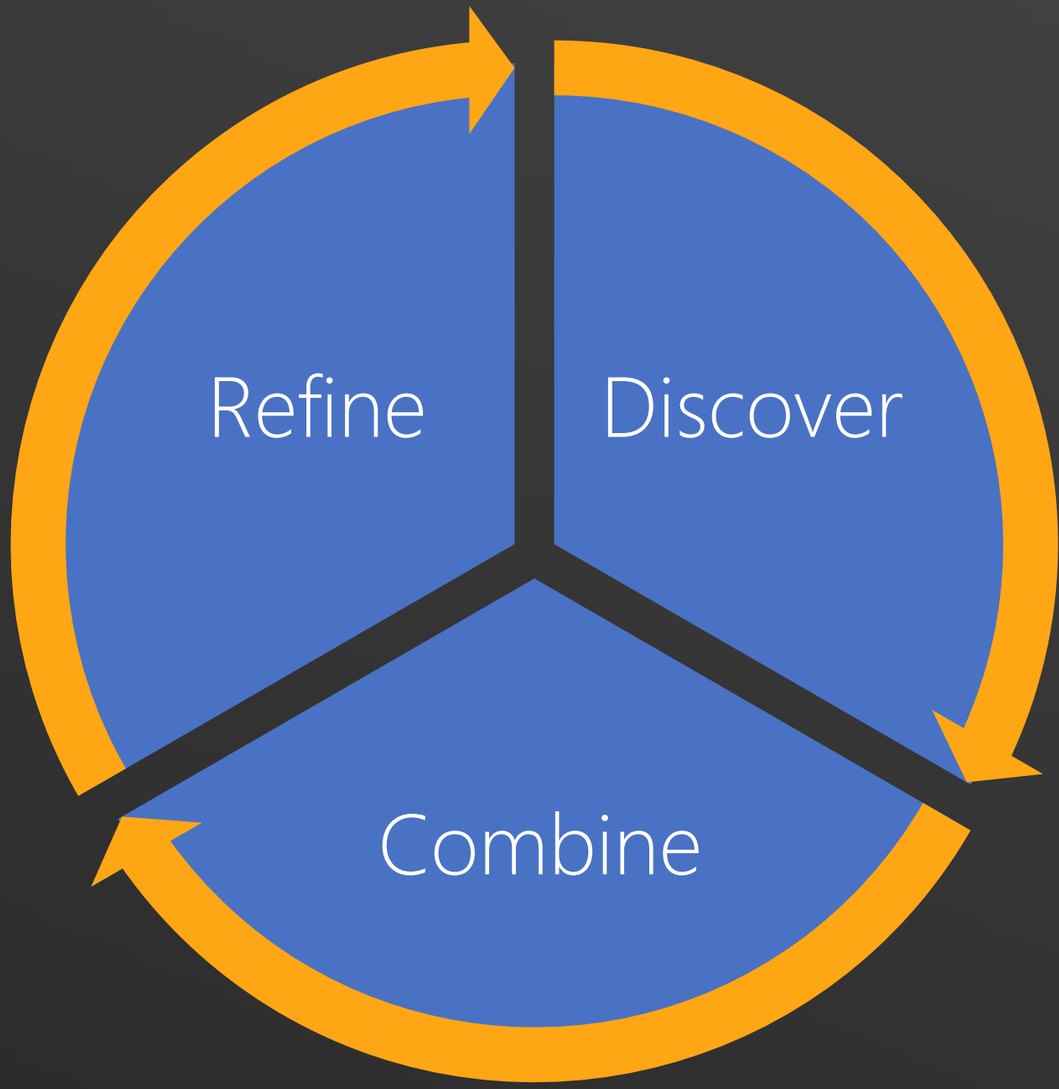
Streaming

Data Movement

THE BIG DATA LIFECYCLE



ENRICH BY CONNECTING TO THE WORLDS DATA



DISCOVER DATA

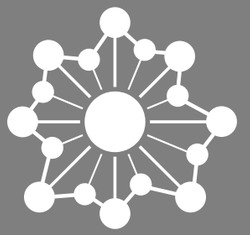
FROM

SEARCH

TO

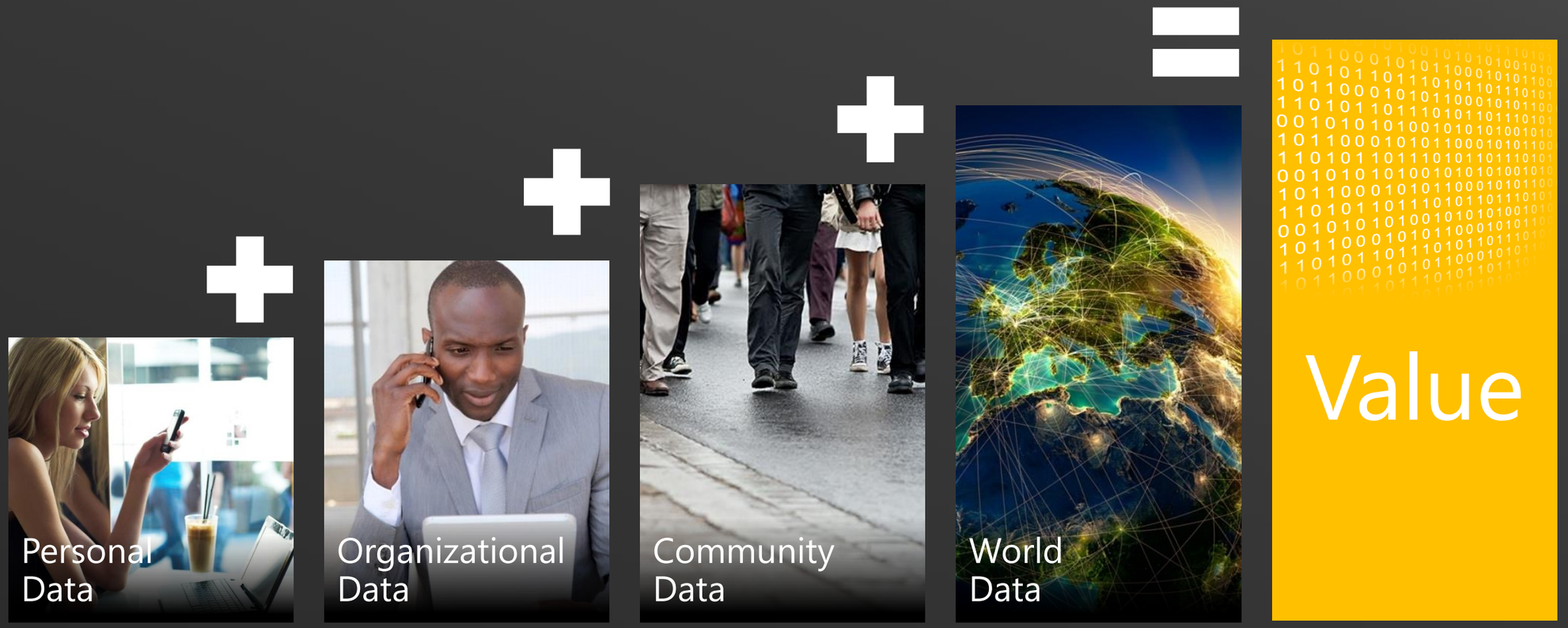
IDENTITY
DOC CONTEXT
SOCIAL GRAPHS

RECOMMEND



DATA EXPLORER
DATA HUB

POWER OF COMBINING THE WORLDS DATA



REFINE DATA

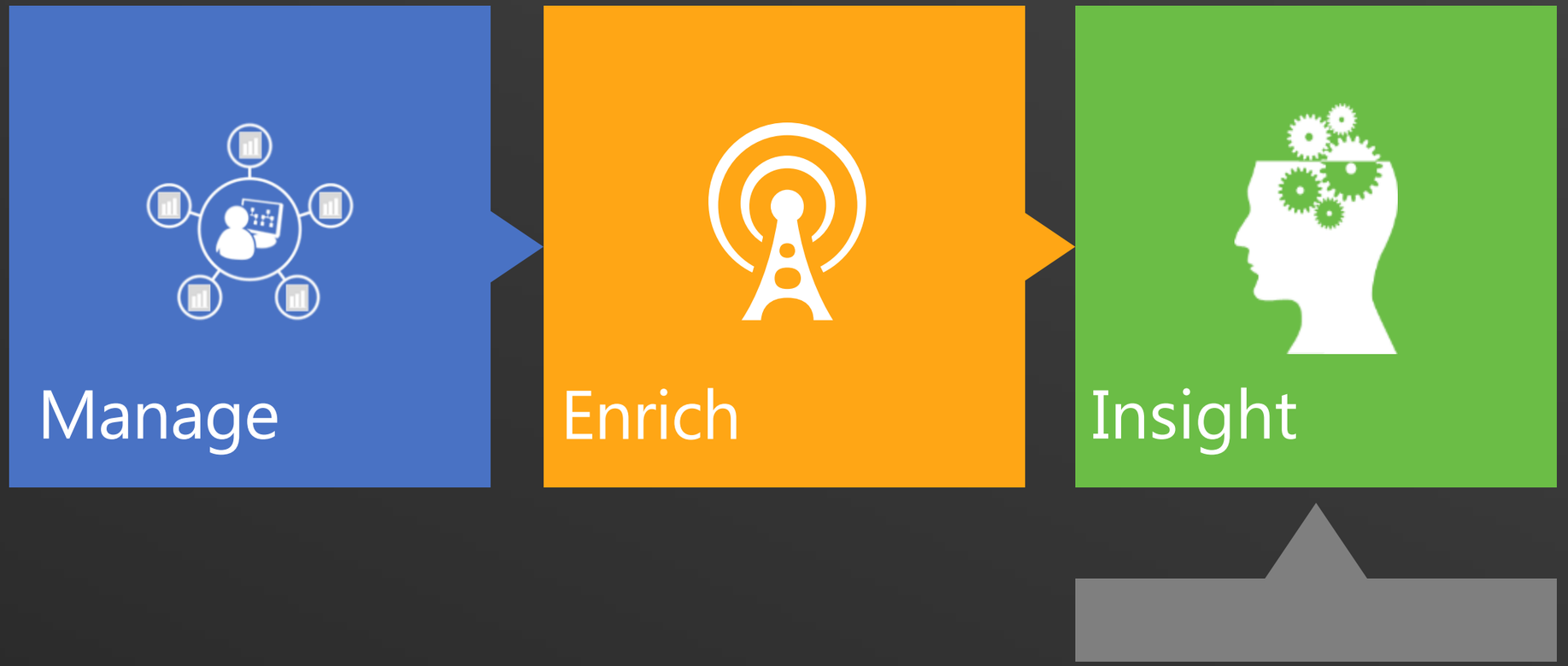
Credible, Consistent Data

Data Mining

Advanced Analytics

Enterprise Information Management & Full Analytic Spectrum

THE BIG DATA LIFECYCLE



INSIGHTS ON ANY DATA, ALL USERS, WHEREVER THEY ARE

INSIGHTS FOR ALL USERS THROUGH FAMILIAR TOOLS

PB

TB

GB

Data Scientists

BI Professionals

Business Analysts

Advanced Analytics from Microsoft and 3rd parties

Self Service Analysis with PowerPivot & Power View

Interactivity & exploration with Hadoop data in Excel

DEMO: FROM DATA TO INSIGHTS!

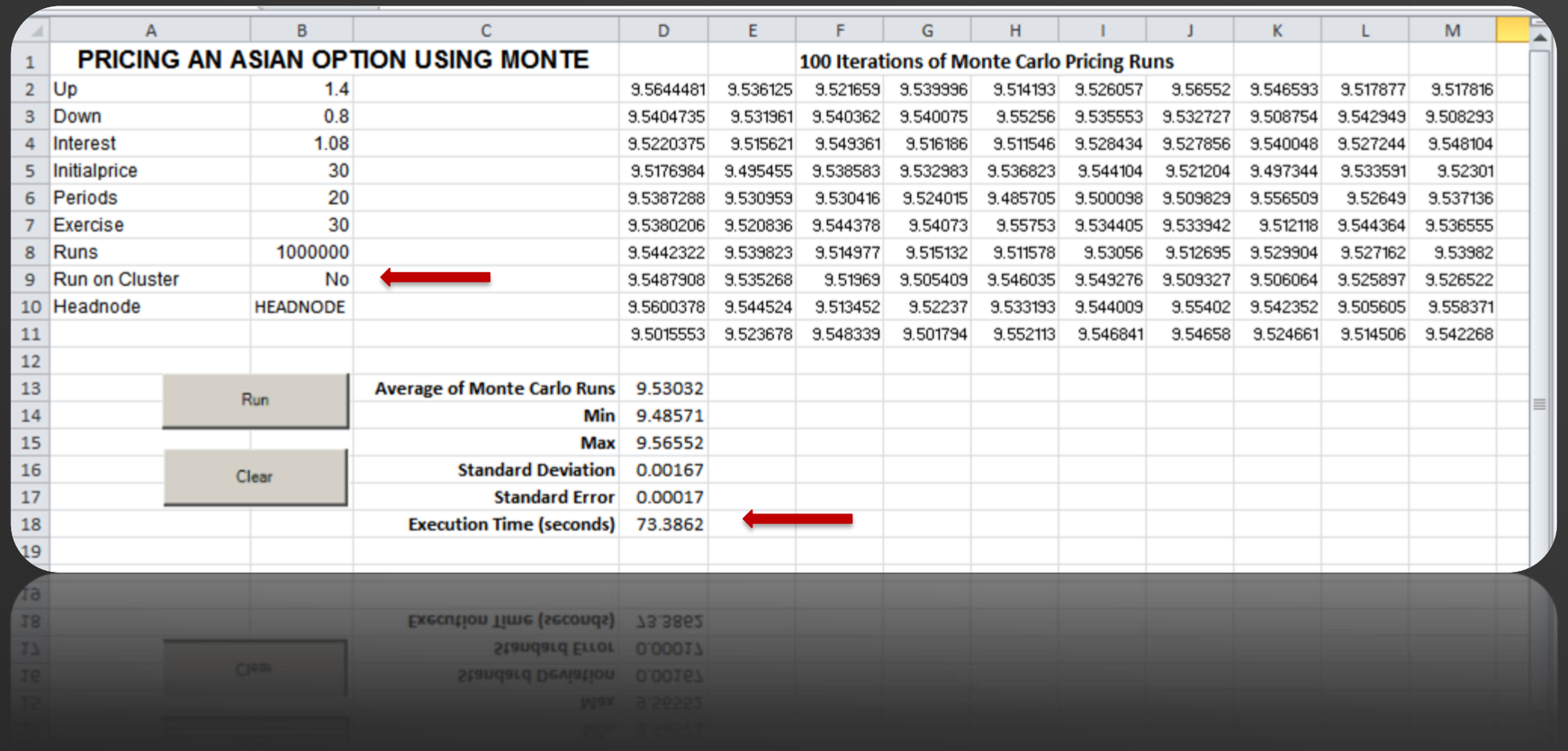
Simplicity

Analysis with familiar tools

Collaboration on insights

MONTE CARLO SIMULATION

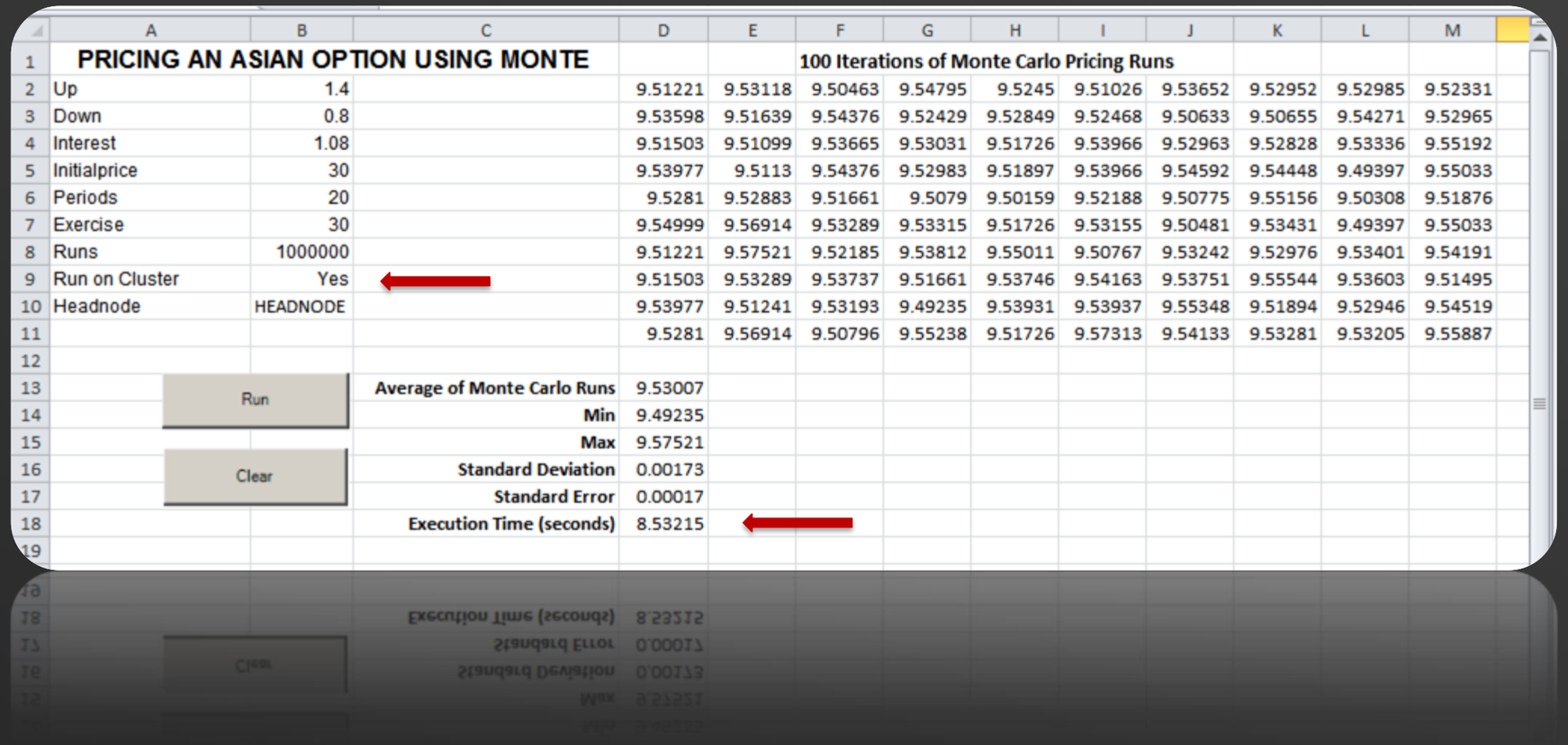
Without a High Performance Computing Cluster



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	PRICING AN ASIAN OPTION USING MONTE					100 Iterations of Monte Carlo Pricing Runs							
2	Up	1.4		9.5644481	9.536125	9.521659	9.539996	9.514193	9.526057	9.56552	9.546593	9.517877	9.517816
3	Down	0.8		9.5404735	9.531961	9.540362	9.540075	9.55256	9.535553	9.532727	9.508754	9.542949	9.508293
4	Interest	1.08		9.5220375	9.515621	9.549361	9.516186	9.511546	9.528434	9.527856	9.540048	9.527244	9.548104
5	Initialprice	30		9.5176984	9.495455	9.538583	9.532983	9.536823	9.544104	9.521204	9.497344	9.533591	9.52301
6	Periods	20		9.5387288	9.530959	9.530416	9.524015	9.485705	9.500098	9.509829	9.556509	9.52649	9.537136
7	Exercise	30		9.5380206	9.520836	9.544378	9.54073	9.55753	9.534405	9.533942	9.512118	9.544364	9.536555
8	Runs	1000000		9.5442322	9.539823	9.514977	9.515132	9.511578	9.53056	9.512695	9.529904	9.527162	9.53982
9	Run on Cluster	No		9.5487908	9.535268	9.51969	9.505409	9.546035	9.549276	9.509327	9.506064	9.525897	9.526522
10	Headnode	HEADNODE		9.5600378	9.544524	9.513452	9.52237	9.533193	9.544009	9.55402	9.542352	9.505605	9.558371
11				9.5015553	9.523678	9.548339	9.501794	9.552113	9.546841	9.54658	9.524661	9.514506	9.542268
12													
13	Run		Average of Monte Carlo Runs	9.53032									
14			Min	9.48571									
15			Max	9.56552									
16	Clear		Standard Deviation	0.00167									
17			Standard Error	0.00017									
18			Execution Time (seconds)	73.3862									
19													

MONTE CARLO SIMULATION

With a 6-Node High Performance Computing Cluster



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	PRICING AN ASIAN OPTION USING MONTE					100 Iterations of Monte Carlo Pricing Runs							
2	Up	1.4		9.51221	9.53118	9.50463	9.54795	9.5245	9.51026	9.53652	9.52952	9.52985	9.52331
3	Down	0.8		9.53598	9.51639	9.54376	9.52429	9.52849	9.52468	9.50633	9.50655	9.54271	9.52965
4	Interest	1.08		9.51503	9.51099	9.53665	9.53031	9.51726	9.53966	9.52963	9.52828	9.53336	9.55192
5	Initialprice	30		9.53977	9.5113	9.54376	9.52983	9.51897	9.53966	9.54592	9.54448	9.49397	9.55033
6	Periods	20		9.5281	9.52883	9.51661	9.5079	9.50159	9.52188	9.50775	9.55156	9.50308	9.51876
7	Exercise	30		9.54999	9.56914	9.53289	9.53315	9.51726	9.53155	9.50481	9.53431	9.49397	9.55033
8	Runs	1000000		9.51221	9.57521	9.52185	9.53812	9.55011	9.50767	9.53242	9.52976	9.53401	9.54191
9	Run on Cluster	Yes		9.51503	9.53289	9.53737	9.51661	9.53746	9.54163	9.53751	9.55544	9.53603	9.51495
10	Headnode	HEADNODE		9.53977	9.51241	9.53193	9.49235	9.53931	9.53937	9.55348	9.51894	9.52946	9.54519
11				9.5281	9.56914	9.50796	9.55238	9.51726	9.57313	9.54133	9.53281	9.53205	9.55887
12													
13	Run		Average of Monte Carlo Runs	9.53007									
14			Min	9.49235									
15			Max	9.57521									
16	Clear		Standard Deviation	0.00173									
17			Standard Error	0.00017									
18			Execution Time (seconds)	8.53215									
19													