

Trasformación digital en *healthcare* para descifrar el genoma en enfermedades raras

De la medicina intuitiva a la empírica

DNA

Laura Escalada Cogollor
Account System Engineer (Laura.escalada@dell.com)

DELLEMC

TGEN & Dell Technologies



TGEN X DELL TECHNOLOGIES

Setting the pace of progress.

Company: TGen | Industry: Genomics Research Institute | Headquarters: Phoenix, Arizona

<https://www.delltechnologies.com/en-us/customer-stories/tgen.htm>

El tratamiento del genoma nos facilita una nueva medicina

❑ Nuevos retos:

- Movimiento de datos ★
- Gestión de datos ★
- Procesamiento de datos ★

❑ Nuevas estrategias en tratamiento de datos ★

- Machine learning
- Deep learning
- Inteligencia artificial

❑ Nuevas necesidades

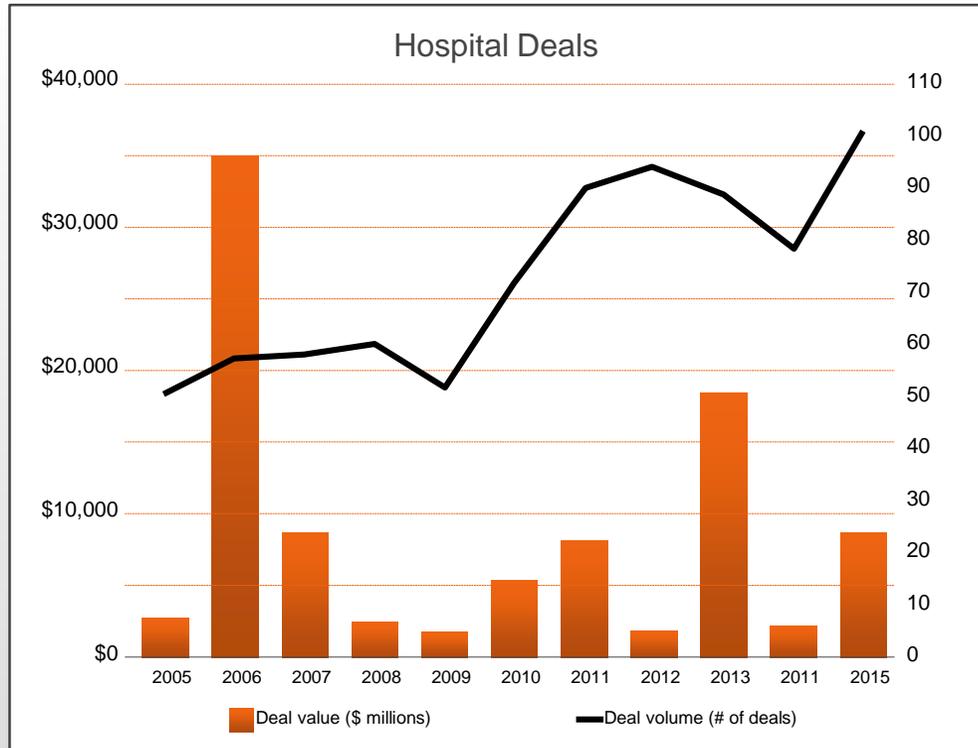
- Necesidad de tecnología
- Necesidad de infraestructura



Comenzamos.....

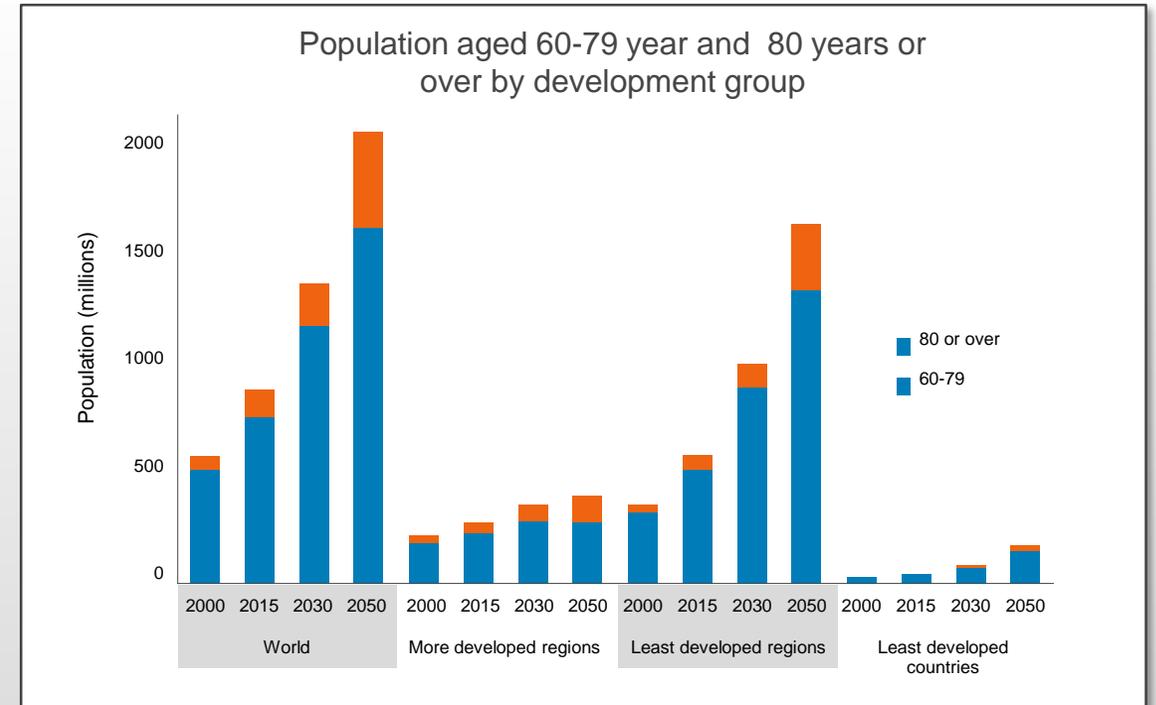


Punto de partida



Source: The Health Care M&A Information

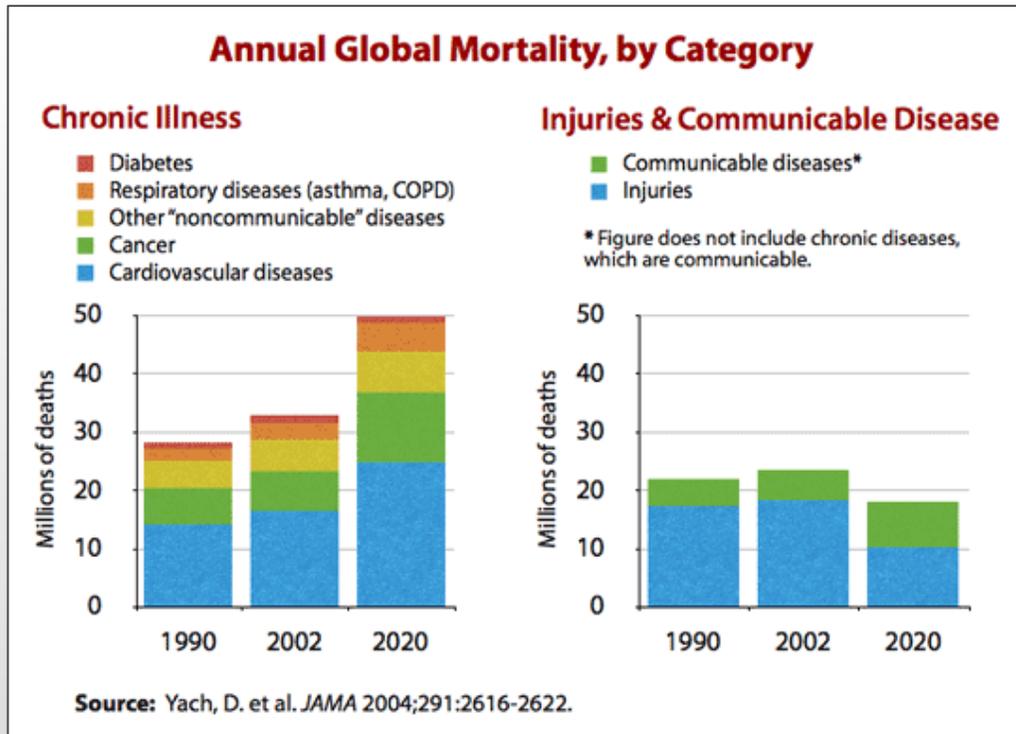
Consolidación de centros



2015 World Population Ageing Report

Envejecimiento de la población

Punto de partida

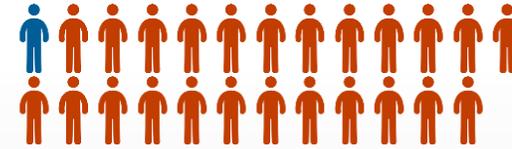


Incremento de Crónicos

Abilify



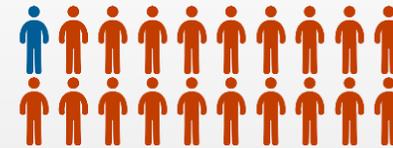
Nexium



Humira

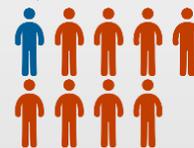


Crestor



Patient helped
 No improvement shown

Cymbalta



Advair Diskus



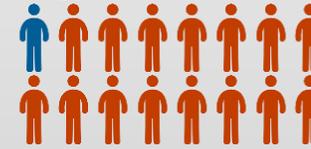
Enbrel



Remicade



Capaxone



Neulasta

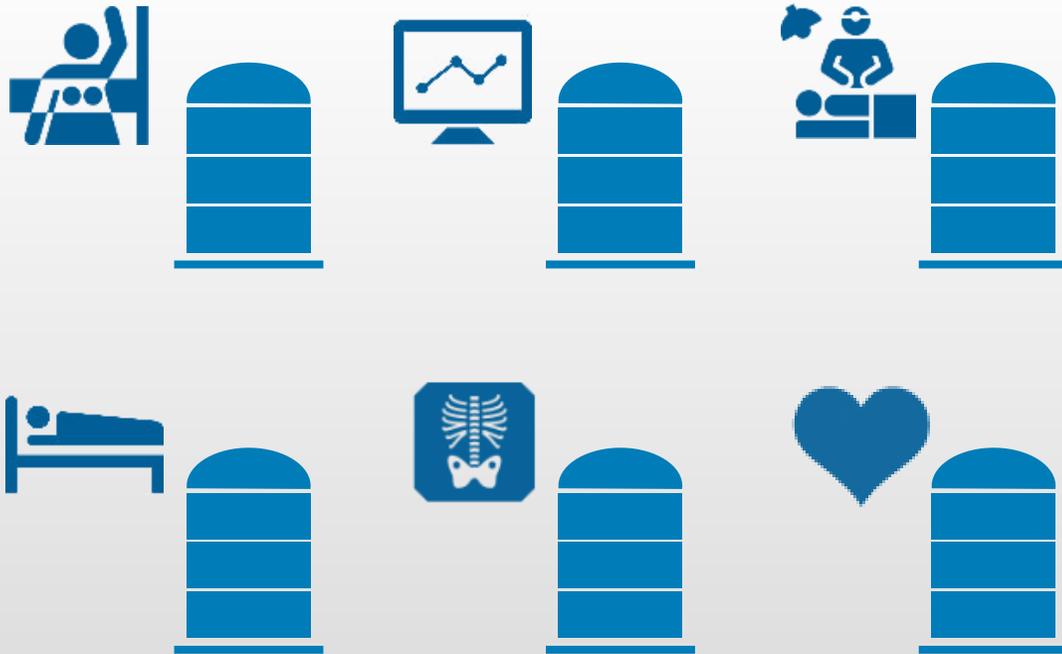


Ten highest grossing drugs in the US fail to improve condition between 3 and 24 people

Bajo ratio de eficiencia

Source: "Schork NJ, "Personalized medicine: Time for one-person trials", Nature 520, 609-611 (30 April 2015)

Nuevas realidades



Incremento de N° de aplicaciones e infraestructura dedicada


Clinical Next Generation Sequencing


Analytics


Digital Pathology

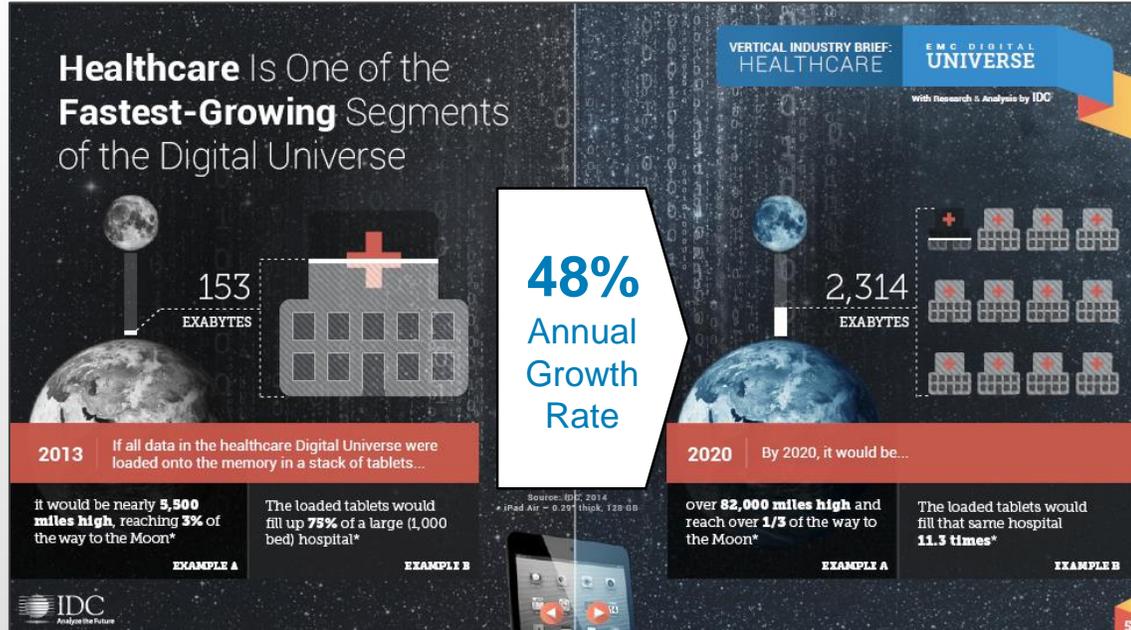

Digital Breast Tomosynthesis


Precision / Personalized Medicine

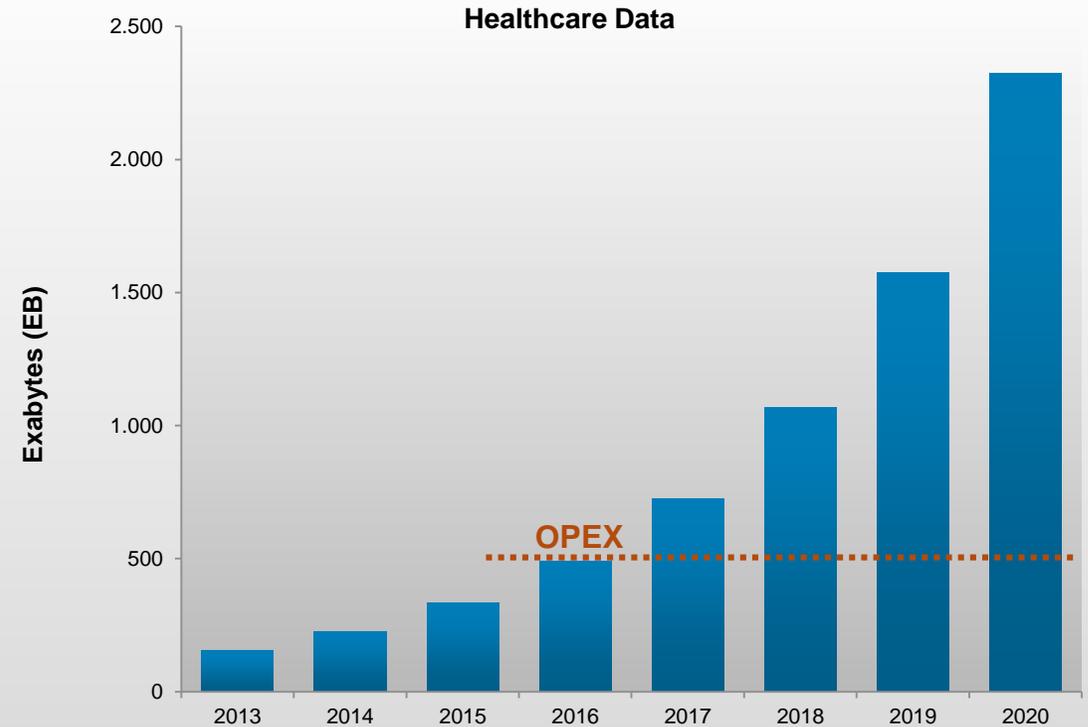

Healthcare Internet Of Things

Digitalización & Nuevos dispositivos Medicina Personalizada

Consecuencias



Crecimiento exponencial de datos



Coste operacional

Consecuencias



Incremento de la colaboración



Cambio de expectativas en los pacientes

Datos del ecosistema sanitario

Aplicaciones Médicas

Tradicionales

Nuevas



PACS / VNA



NGS Clínica



EHR



Digital Pathology



Ecografías



Tomografías



Estudios del sueño

Servicios generales de IT



Sistemas corporativos



Contenido web



Video vigilancia y grabaciones



Archivos médicos



Entornos de colaboración



Otros casos de uso



Monitorización y diagnóstico remoto



Ambulancias inteligentes

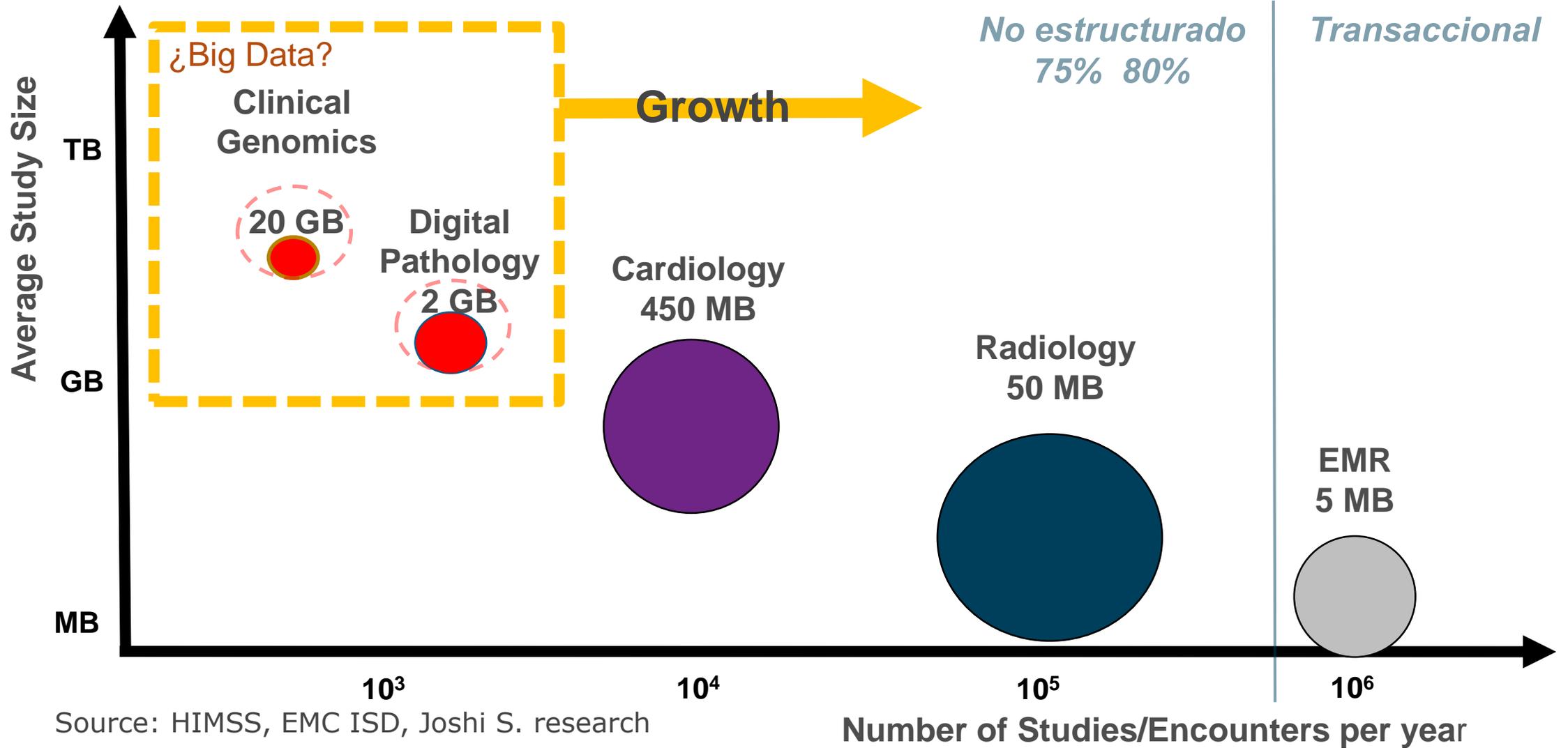


IoT



Nuevas analíticas (P4)

Nuevos Datos



Nuevos límites

Performance ↑



Analytics

Compute: CPU
Throughput: MB/s
Concurrency: 10s
Scale: GBs–TBs



Machine Learning

Compute: CPU & GPU
Throughput: GB/s
Concurrency: 100s–Ks
Scale: 10s TB's–PBs



Deep Learning

Compute: 10s–1000s GPU
Throughput: 10s–100s GB/s
Concurrency: 1000s–Ms
Scale: 100s TBs–10s PBs



Scale →

¿Cómo integramos el BD en el entorno Sanitario?

Aplicaciones Médicas

Tradicionales

Nuevas



PACS / VNA



NGS Clínica



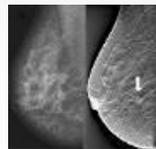
EHR



Digital Pathology



Ecografías



Tomografías



Estudios del sueño

Servicios generales de IT



Sistemas corporativos



Contenido web



Video vigilancia y grabaciones



Archivos médicos



Entornos de colaboración



Otros casos de uso



Monitorización y diagnóstico remoto



Ambulancias inteligentes



IoT



Nuevas analíticas (P4)

Arquitectura tradicional



Arquitectura Tradicional

NGS SERVICE PROVIDER

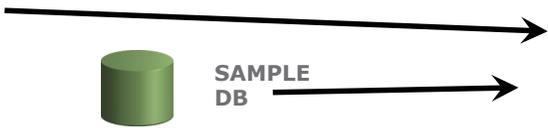


ASPERA



Copia

SMB



Copia



Copia



Copia



Copia



S3



SAMPLE DB

CLINICAL DB

ANNOTATION DB

PACs DB



Cloud Pública



Colaboradores

Cloud Privada



Real Time

IoT - Monitorización Remota



Continuidad De servicio



Usuarios



SMB

NFS

HDFS

Cuales son los retos de implementar “Big Data”



Tiempo y recursos para provisionar los nuevos entornos



Sin flexibilidad de herramienta de análisis



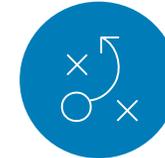
Complejidad al añadir casos de uso a los entornos existentes



Infraestructura dedicada
Ineficiencia en el uso de los recursos



Complejidad en la securización del dato



Coste de adquisición y de mantenimiento

Y si además tenemos en cuenta que...

“Necesitamos entornos de test y desarrollo para no impactar a producción”

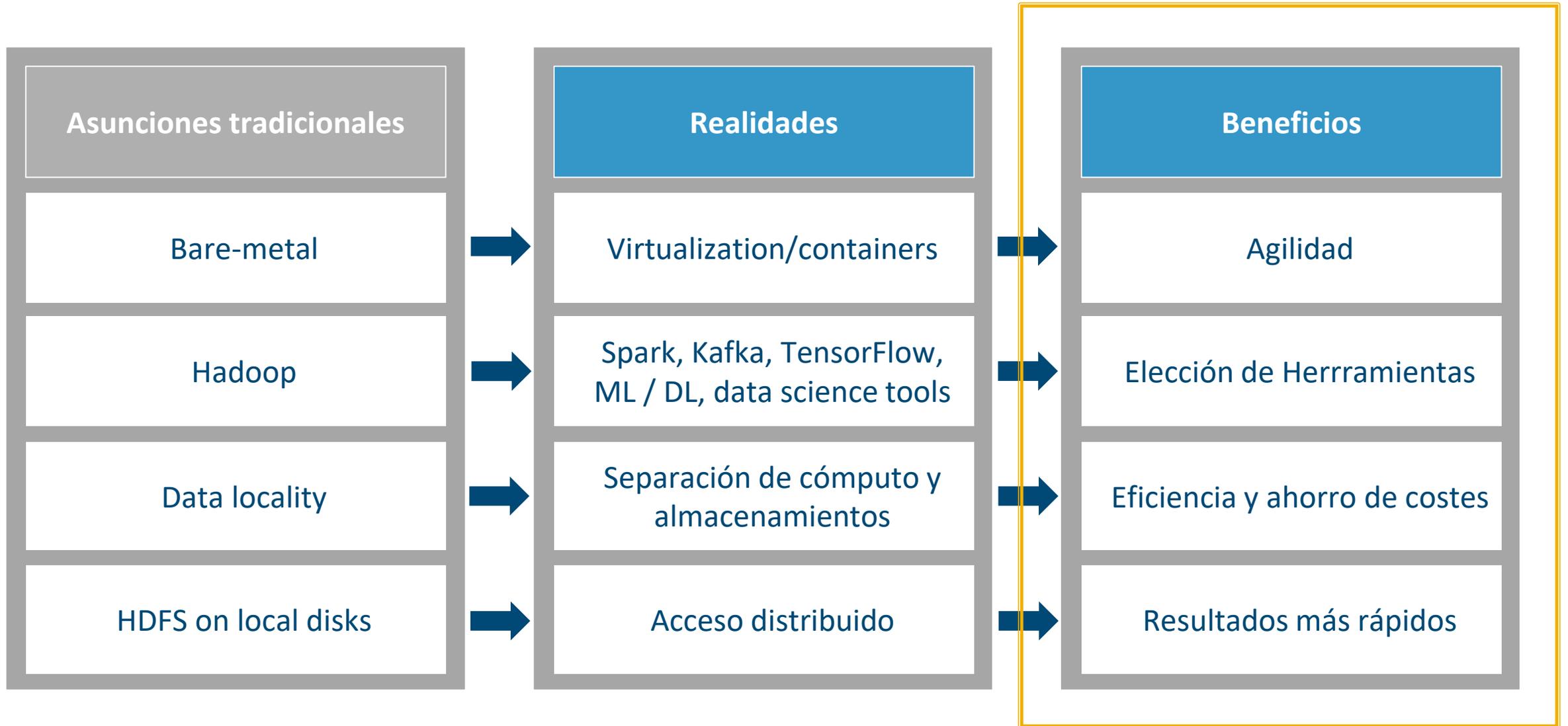
“Necesitamos escalar computación y almacenamiento de manera independiente”

“Necesitamos gestionar la seguridad por entorno, usuario y administrador”

“Necesitamos agilidad de despliegue tipo Cloud”



Realidades del Big Data....y beneficios

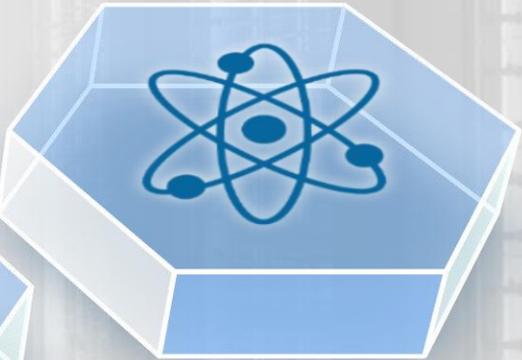
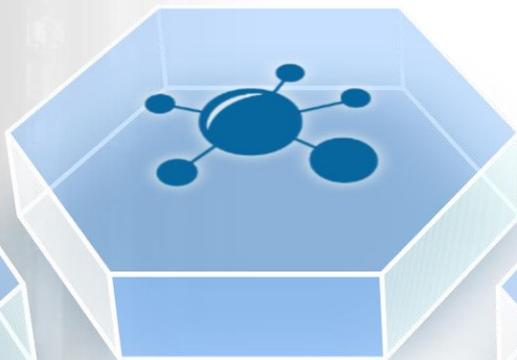
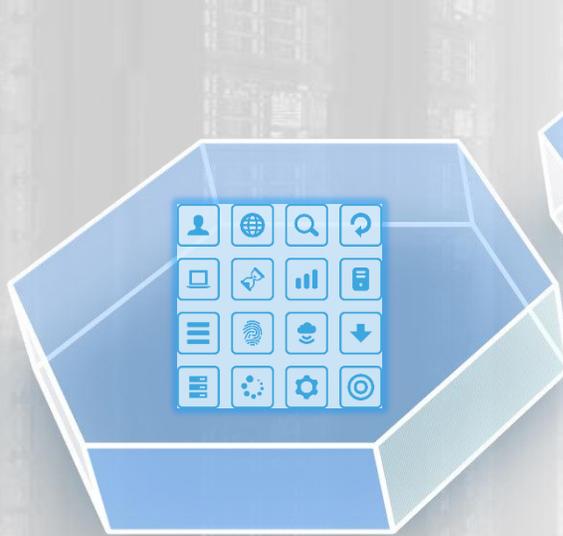


**Necesitamos un
verdadero DataLake**



Aspectos a tener en cuenta

Entornos Tradicionales
y de nueva generación



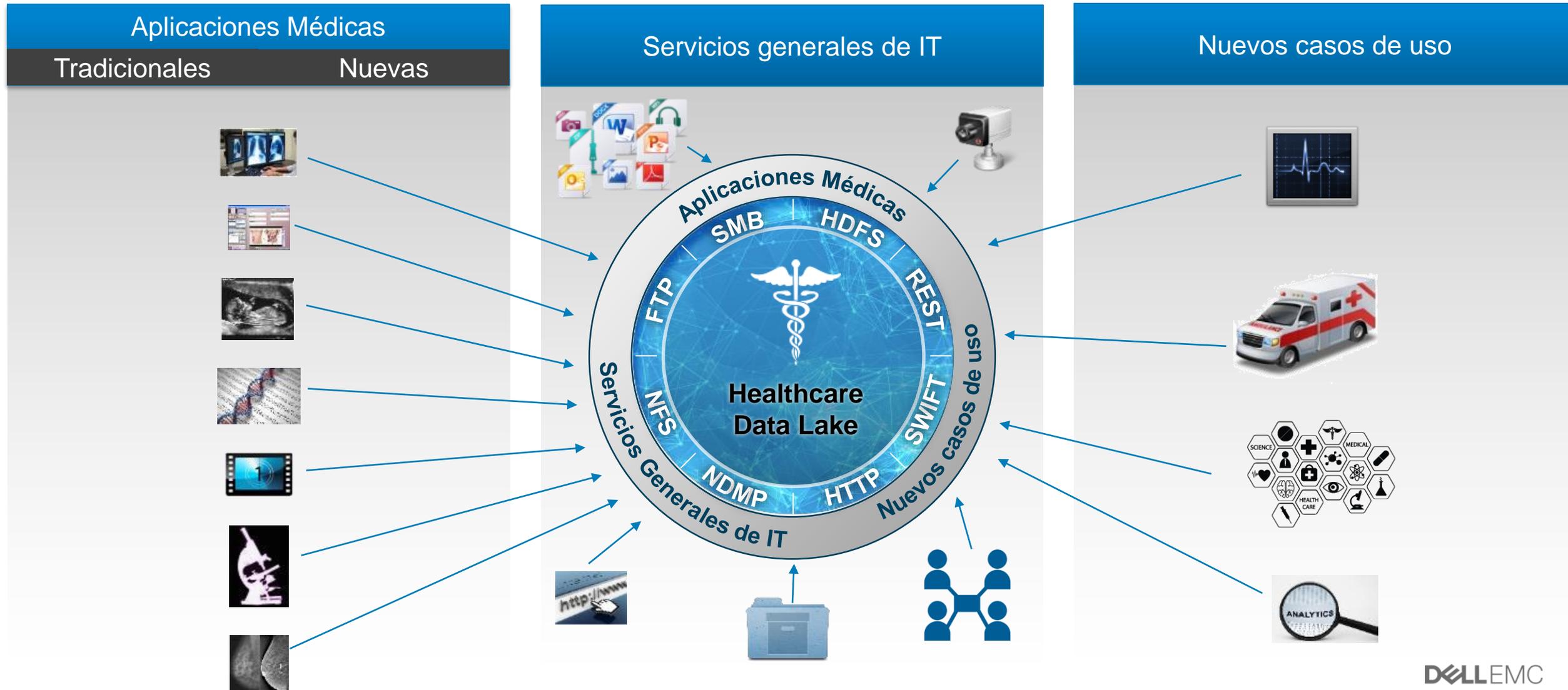
El análisis es un
proceso continuo

Magnitudes nunca antes vistas han
cambiado los principios de diseño

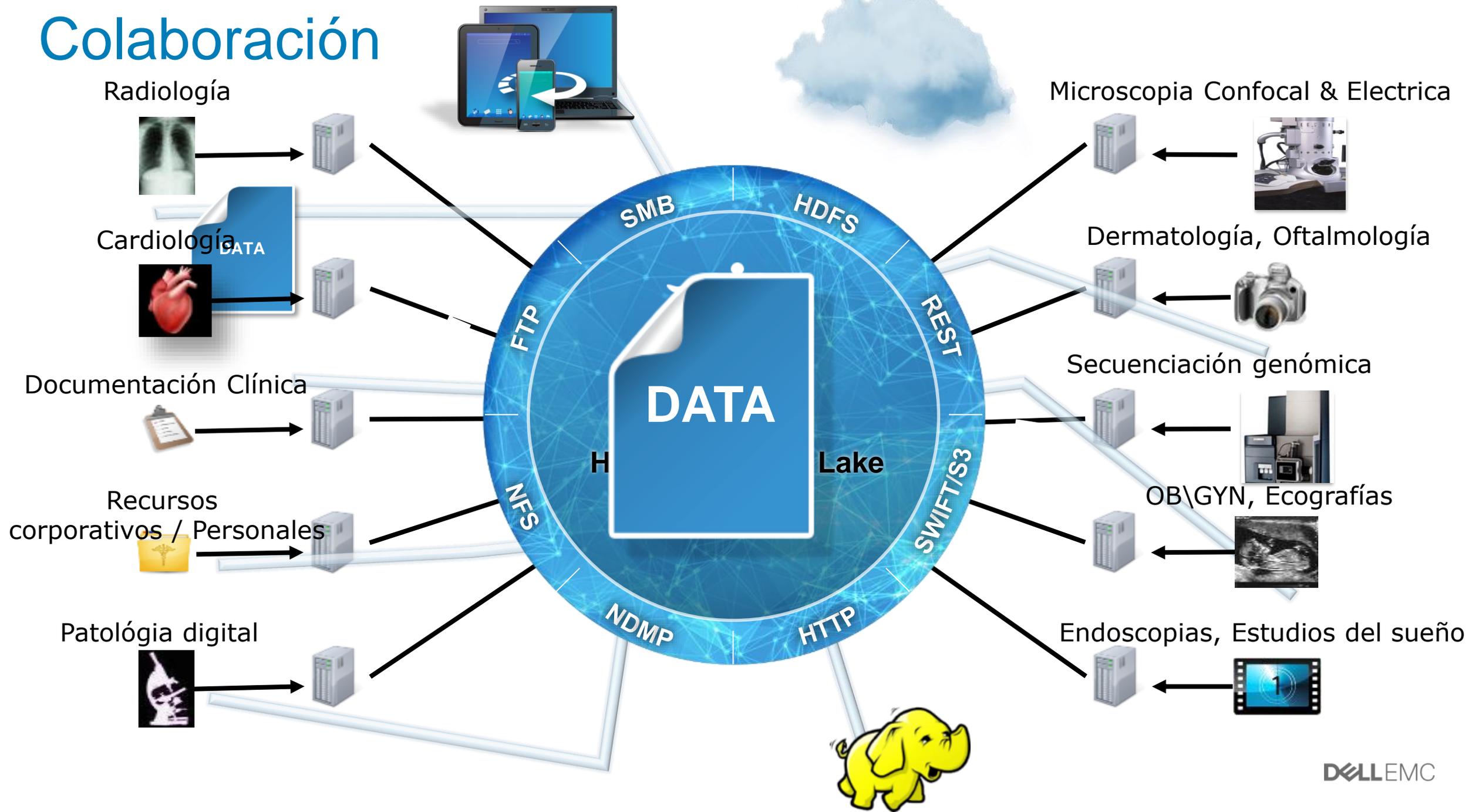
La agilidad y
la colaboración es crítica

Healthcare Datalake - Un Sistema para presente y futuro

Usuarios, datos y aplicaciones



Colaboración



Un único Data Lake – IT as a Service

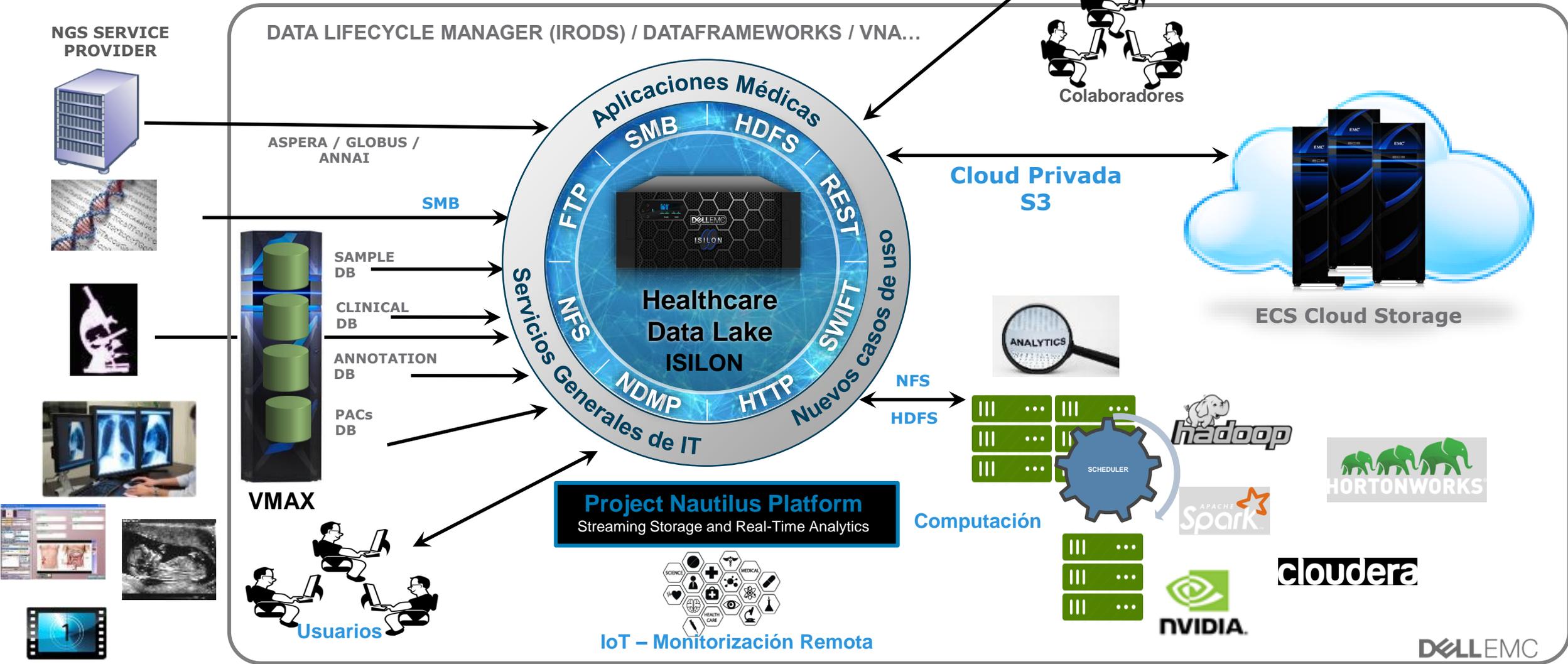


- **Permitimos escalar en capacidad, dispositivos, aplicaciones, modalidades, servicios**
 - Hasta 100PB, 15TB/s....
 - Sin limitaciones en capacidad, rendimiento, etc
- **Simplificamos la administración - Eliminamos multiples repositorios**
 - Siempre un file system
 - Consolidamos workflows clínicos y no-clínicos
- **Todo funciona por políticas - Alineamos el IT al servicio, al negocio**
 - Las funcionalidades y la protección se puede configurar a nivel de carpeta
- **Facilitamos el intercambio de información - Colaboración**
 - No importa como llega el dato, importa el dato
- **Creamos servicios a un coste justo**
 - Desde flash hasta la nube, transparente para la aplicación
 - Control y previsión de costes y consumos, planificación de crecimientos
- **Integramos las nuevas demandas (ej: genómica, patología digital, etc)**
 - Arquitecturas de referencia con los nuevos casos de uso
- **Aceleramos los procesos de analíticas y eliminamos duplicados**
 - Nativamente integrado con herramientas de análisis de Big Data
- **Garantizamos la preservación de la información**
 - Auditoría de usuario y administración, Periodos de Retención, WORM
 - Blindado ante refrescos tecnológicos

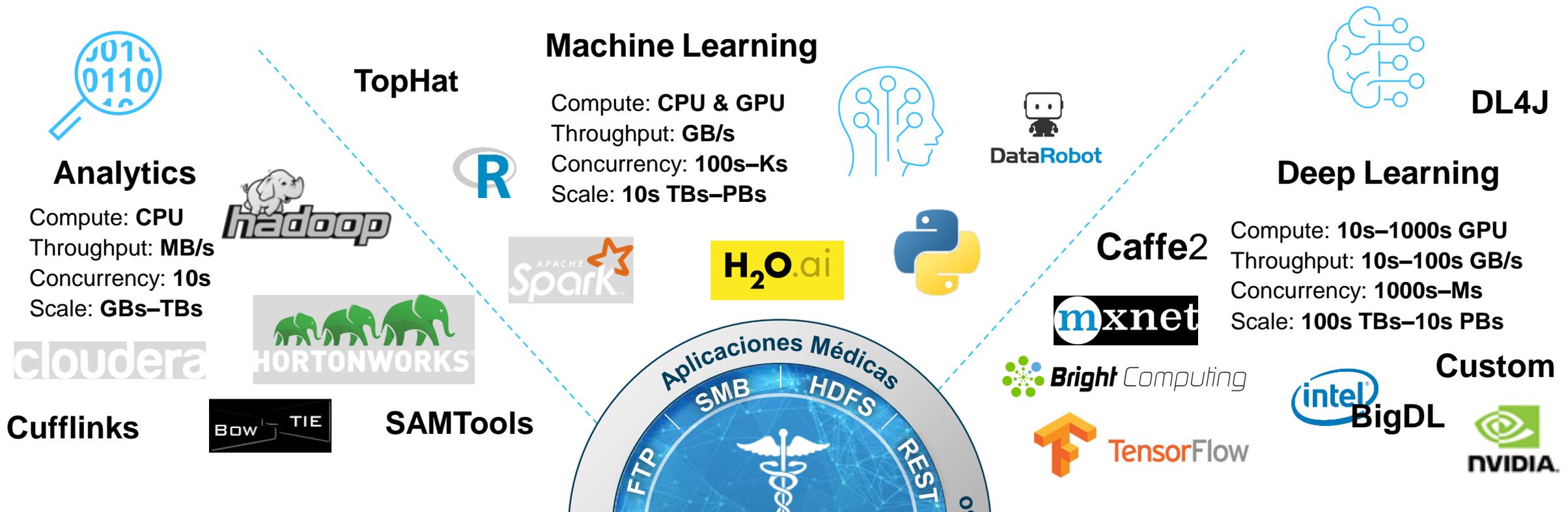


Arquitectura global

Arquitectura Global



Integrando Big Data as a Service en el Datalake



- ✓ Minimiza el movimiento de datos
- ✓ Permite elegir la herramienta del análisis
- ✓ Sencillo y flexible en el despliegue de entornos

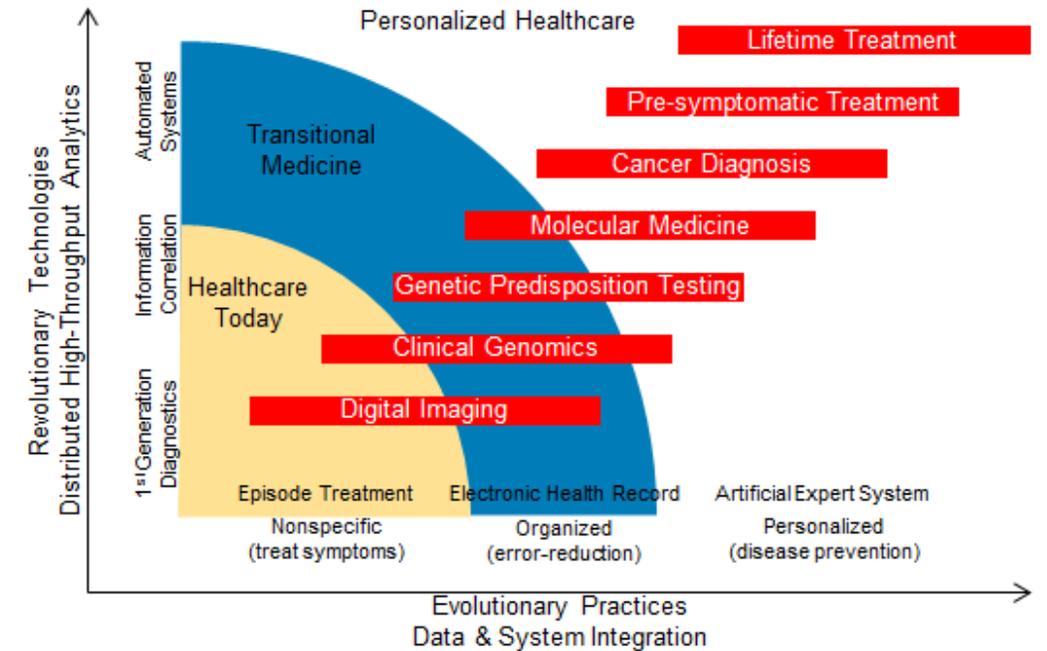
- ✓ Escala independiente casos de uso
- ✓ Permite múltiples topologías
- ✓ Punto único de soporte



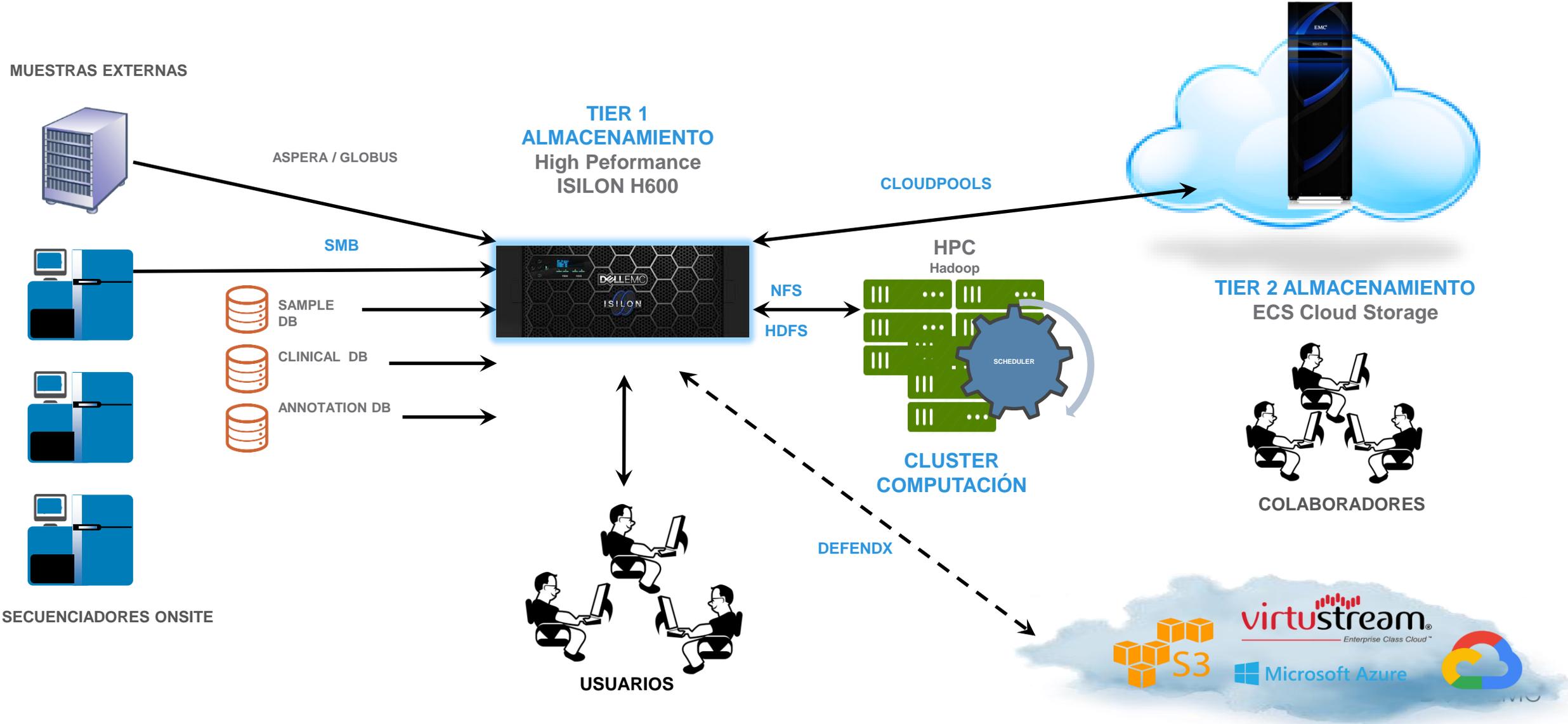
Arquitectura NGS

Secuenciación genómica

- Perfil del Ciclo de Vida de las muestras
 1. Secuenciación de las muestras
 2. Proceso de las muestras y generación de archivos de resultados
 3. Archivo y protección de ficheros de muestras y resultados
 4. Reanálisis de las muestras en función de diferentes casos de uso
- Tipos de dato – Exoma (22GB) y Panel (5GB)
- Tipos de usuario o consumidores del dato – Acceso multidisciplinar
 - Secuenciador (típicamente NFS)
 - Nodos de cabecera y computación (típicamente NFS)
 - Investigadores (FTP/NFS/SMB)
- Nivel de Acceso
 - Tipo A (“Productivo”).- Muestras del último año y potencialmente al procesado/resumen de muestras históricas
 - Tipo B (“Archivo/Preservación”).- Datos completos de las muestras históricas (>1 año)
- Retención/Custodia temporal de los datos
 - Corto – 1 Año – Máxima utilización
 - Medio – 5 Años – Ley
 - Largo – Ilimitado – Histórico (Renálisis, estadísticas)



ARQUITECTURA REFERENCIA NGS + ML



Referencias



Referencias en Investigación y Sanidad



En resumen...



Conclusiones

En la transformación digital para descifrar el genoma humano, son **fundamentales**:

1. Algoritmos de **secuenciación**
2. **Supercomputación** / alta capacidad de procesamiento
3. Los **datos!!**
 - Repositorios comunes, grandes y compartidos.

Y sobre todo, la implantación de **arquitecturas de referencia**, consolidadas y flexibles para:

- todos los entornos el ecosistema salud (sanitario e investigación) y
- herramientas del presente y del futuro.



Muchas gracias por su atención!!