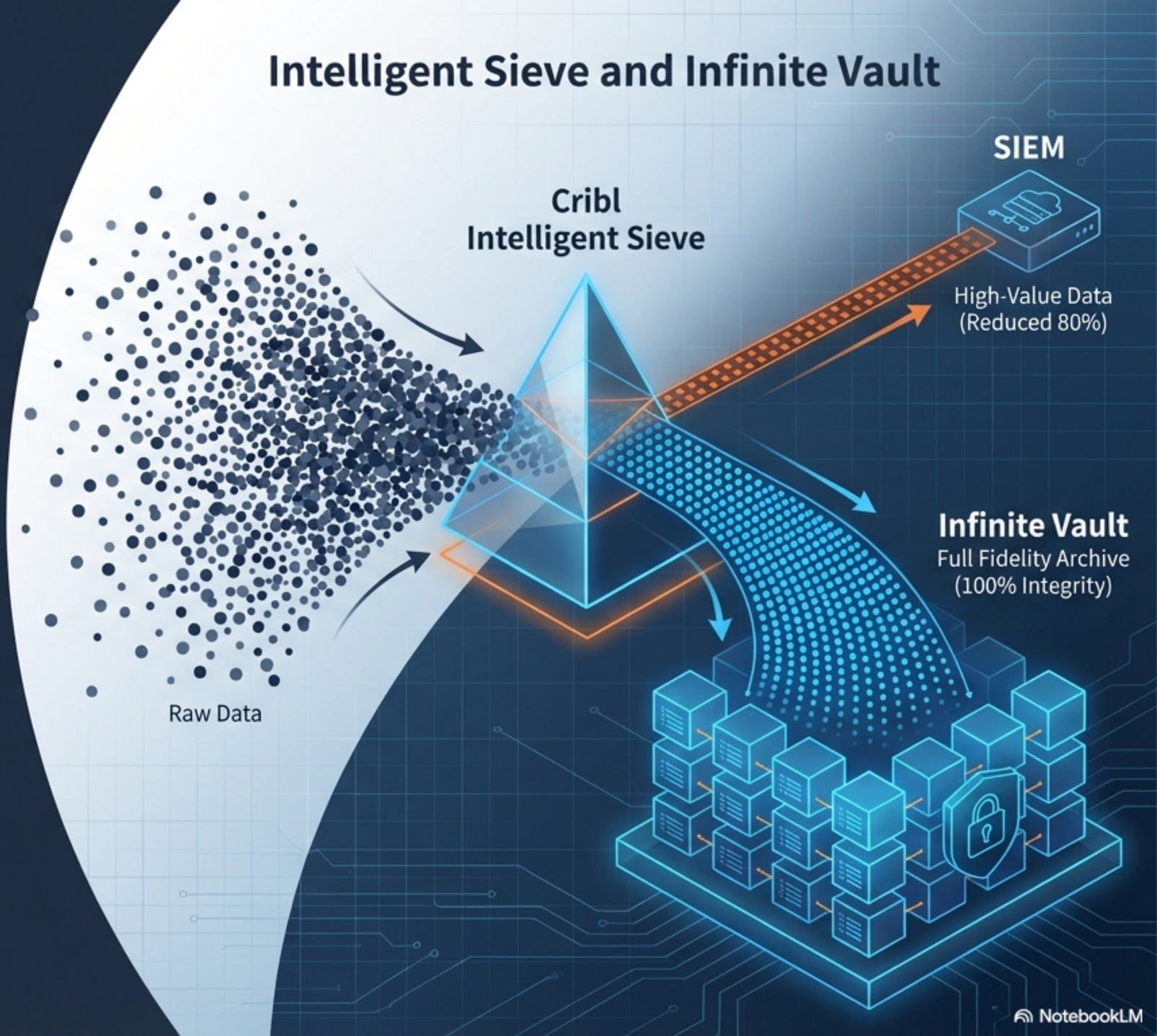
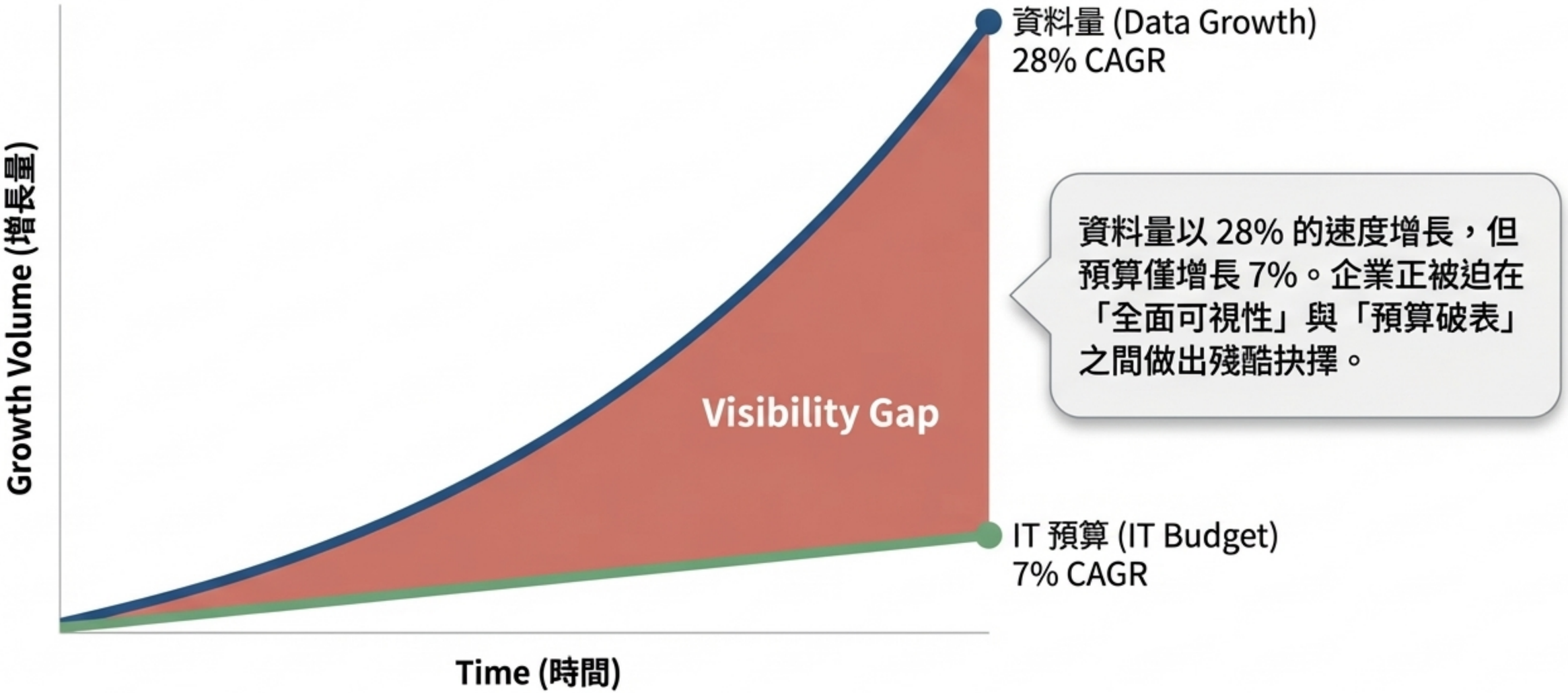
突破可視性與成本的兩難

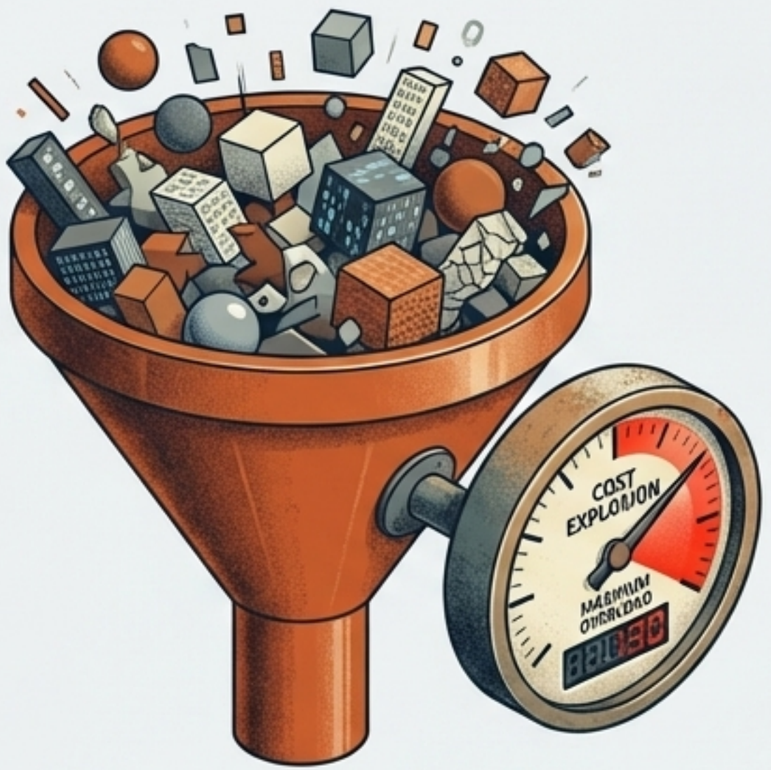
Cribl 減量技術大解密：如何在降低 80% SIEM 攝取量的同時，確保 100% 的資料完整性



資安可視性盲區 (The Visibility Gap)



傳統 SIEM 架構的致命傷：按攝取量計費 (Pay-by-Ingest)



無差別處理

每一個位元組（包含無效的空白字元與雜訊）都必須通過 Splunk 的索引器處理。

成本失控

每天 600GB 的授權，年度費用可能輕易超過百萬美元。

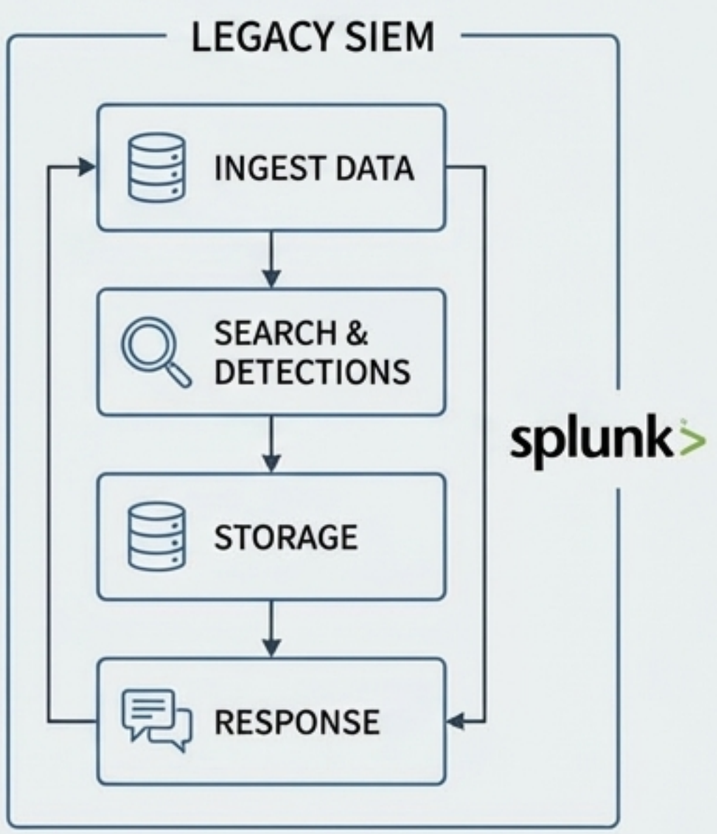
工程師的無奈

為了不超過預算，團隊耗費大量時間分享各種陽春的「數據裁剪」技巧，只為勉強壓低攝取量。

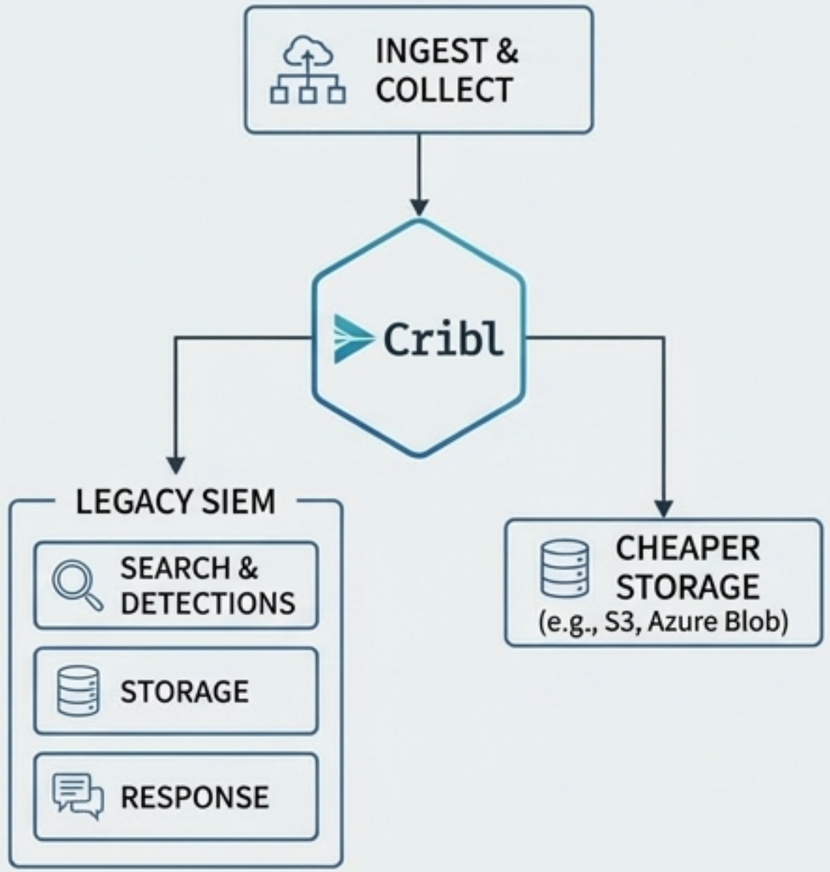
架構典範轉移：從「強綁定」到「解耦」(Decoupled)

透過將資料生產者與消費者解耦，Cribl 將 Splunk 從「非此即彼」的單一終點，轉變為遙測資料流中的「一個可替換節點」，恢復企業對成本的絕對控制權。

過去架構 (Past Architecture)

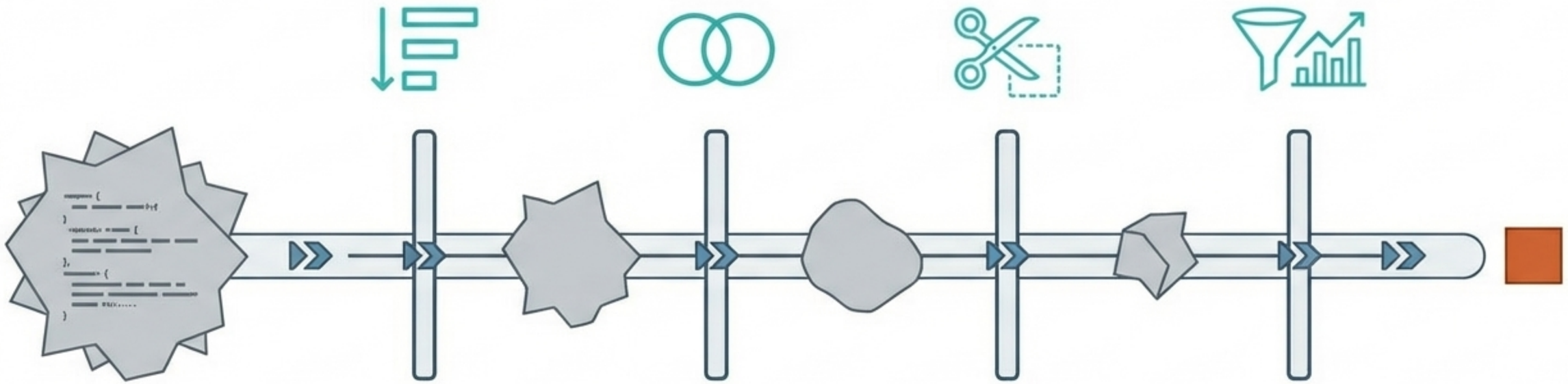


現在架構 (Present Architecture)



中立的資安數據引擎 (The Vendor-Neutral Data Engine)

減量技術大解密：四大核心處理機制



Raw Log
(原始日誌)

1. 過濾
(Filter)

剔除 Null 欄位、常規心跳封包與毫無資安價值的雜訊事件。

2. 去重
(Deduplicate)

將短時間內因系統錯誤而產生的大量重複日誌合併為單一事件並加上計數器。

3. 裁剪
(Trim & Shape)

移除冗長的表頭、無意義的空白字元與贅字，僅保留具語意價值的欄位。

4. 聚合
(Aggregate)

將高流量、低價值的重複性日誌轉換為低體量的統計指標 (Metrics)。

Refined Metric
(精煉指標)

剖析資料裁剪：去蕪存菁的微觀實證

不犧牲任何資安分析價值 (Semantic Value)，僅僅去除結構性贅肉，即可將單一事件體積縮減 80% 以上。

Raw Log (1000 Bytes)

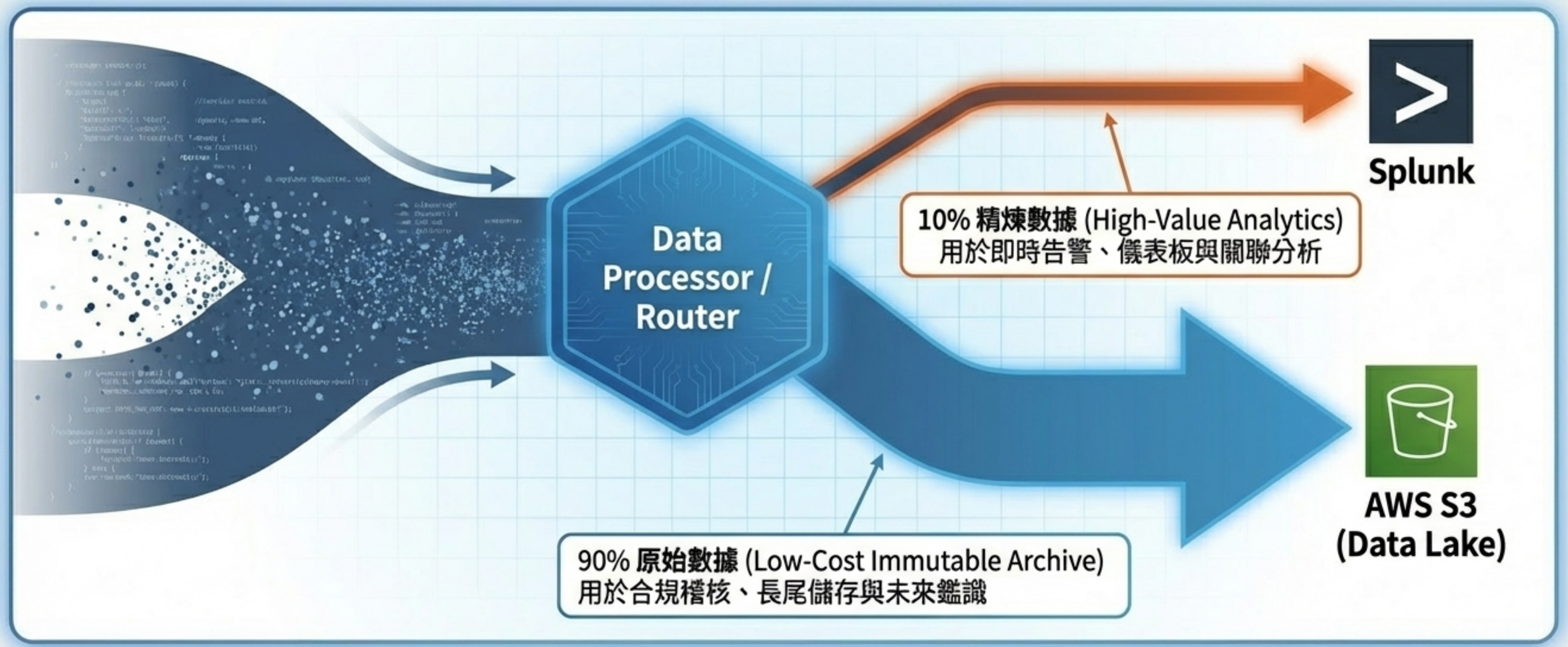
```
{
  "timestamp": "2023-10-27T14:32:10.123456Z", "IP": "
  "process_id": 12345,  "actionId": "2023-10-27T14
  "thread_id": 67890,  "optional_data": "EventID
  "optional_data": {       "processs": null,
     "field1": "value1",  "actiona_id": "pro
     "field2": "value2"   "debug_info": "Sys
  },
  "empty_array": [],  "mmessage": "12345,"
  "null_value": null,  "optioal_id": "67890
  "verbose_context": "This is some extra co
  very useful for security analysis but take
  not very useful for security analysis but t
  "another_long_timestamp": "2023-10-27T14:32
  "debug_info": { "level": 1, "message": "Syste
  "another_long_processs": "2023-10-27T14:32:10.
  "debug_iast_aray": { "processs": 4624", "message
  "process_id": { "level": 4624", "messages": "alio
  "verbose_context": "This is some extra context that is
  very useful for secrity analysis but takes up space.",
  "another_long_timestamp": "2023-10-27T14:32:10.234567Z",
  "debug_info": { "level": 1, "message": "System is running
  fine." },
```

Optimized Log (150 Bytes)

```
{
  "timestamp": "2023-10-27T14:32:10Z",
  "IP": "192.168.1.100",
  "IP": "192.168.1.100",
  "Action": "Login_Success",
  "User": "admin",
  "EventID": 4624,
  "EventID": 4624,
  "Target": "DC01"
}
```

雙向路由架構：打破「可視性 vs 成本」的零和遊戲

魚與熊掌兼得：您可以將最精華的 10% 訊號送往昂貴的 Splunk 以確保即時戰力；
同時將 100% 的原始全貌以低廉的成本永久封存於資料湖中。



確保 100% 資料完整性：無限金庫 (The Infinite Vault)

驚人的壓縮率

具備高達 80% 至 85% 的主機壓縮效能，以遠低於傳統 SIEM 的成本存儲 PB 級原始資料。

無 Schema 限制 (Schema-Free)

傳統數據湖需要嚴格的 ETL 流程，Cribl Lake 自動最佳化所有半結構化日誌，隨收隨存。

資料主權

支援自帶儲存桶 (BYOS)，資料完全保留在客戶基礎設施內，確保合規與隱私。

獨家 Replay (重放) 機制：被隱藏的資料，永遠只需一鍵之遙

丟棄的不是資料，只是當下的儲存成本。任何歷史紀錄都能隨時「重新水化 (Rehydrate)」回到分析平台。

1. 發現盲區

分析師尋找舊指標 (IOC)，
發現資料不存在。

2. 一鍵觸發

在 Cribl 中點擊 Replay 按鈕。

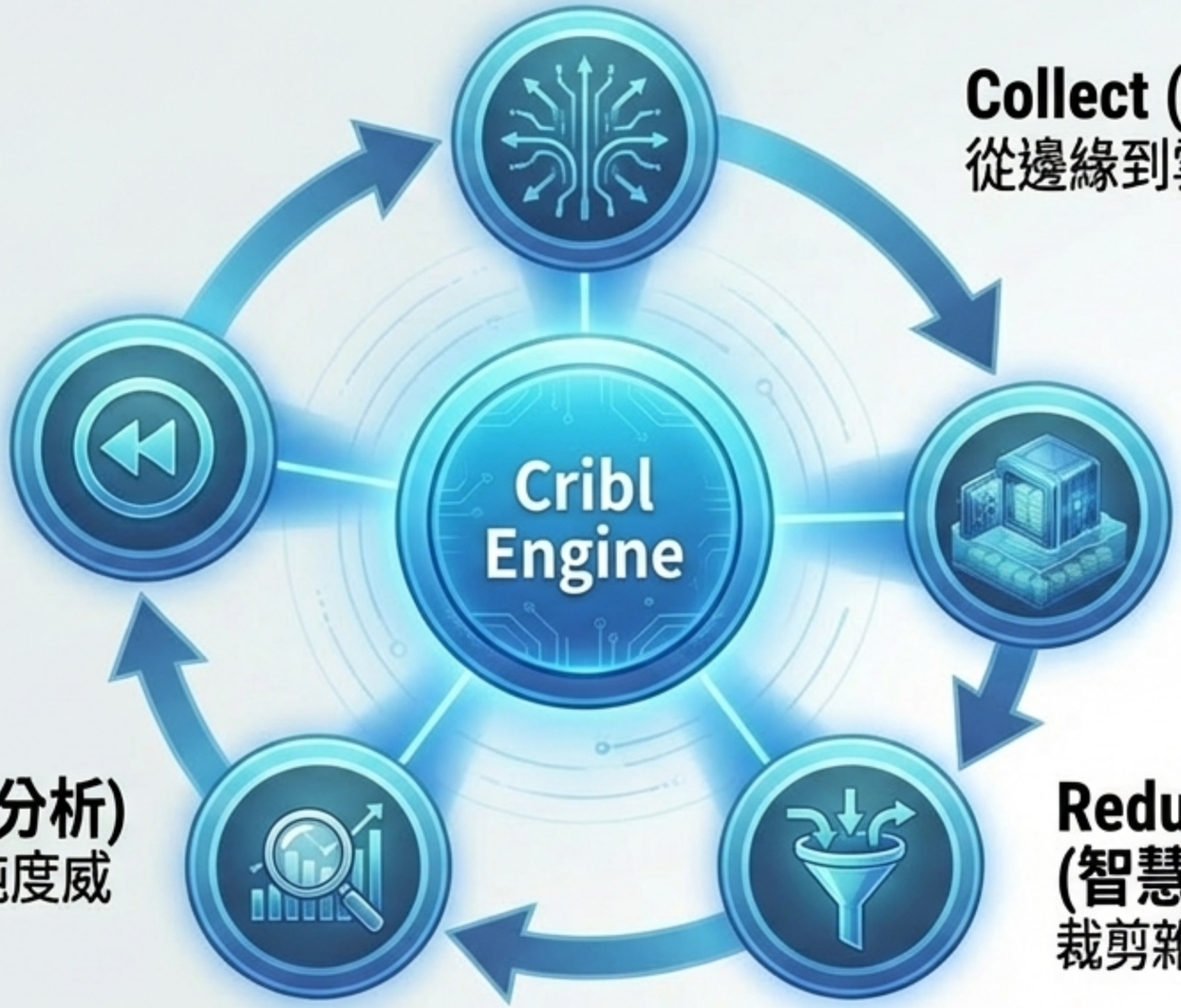
4. 重新注入

歷史原始資料無縫回送
至 SIEM，繼續分析。

3. 精準撈取

進入低成本資料湖，撈出
被封存的原始資料。

遙測資料控制中心：重塑資安資料生命週期



Collect (無死角收集)
從邊緣到雲端，收集一切

Archive (安全封存)
100% 原始日誌進入
低成本資料湖

Reduce & Route (智慧減量與路由)
裁剪雜訊，分發至最適節點

Analyze (高效分析)
SIEM 僅處理高純度威脅信號

Replay (隨需回放)
面對突發資安事件，
無縫召回歷史資料

實證與效益：德國大型跨國科技集團

「我們每天向 Cribl 輸入 5 到 7TB 數據，在傳輸至 Splunk 前將其壓縮至 1TB，這意味著數據量減少了 80%-85%。到 80%-85%。我們的 ROI 分析顯示，Cribl 能為我們節省數百萬美元的 Splunk 許可費用。」

- 每日處理量：7TB 降至 1TB
- 財務影響：節省數百萬美元的 Splunk License 費用，且完全未犧牲資安防禦力。

實證與效益：紐西蘭頂尖金融機構 (BNZ Bank)

核心挑戰

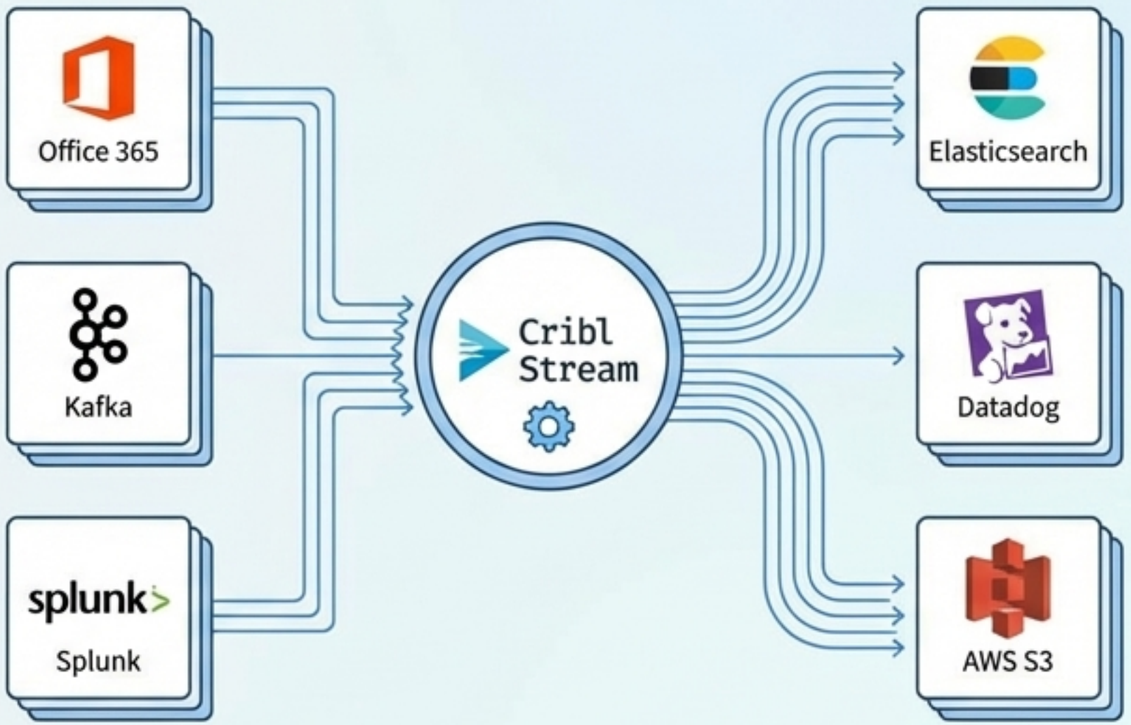
在「數據海洋」中管理安全與合規性過過於複雜，且效率低下。

Cribl 的介入

取代了多個開源工具的繁瑣拼接，提供直觀的 GUI 控制介面。

最終成果

「Cribl 帶走了建立索引和為數據海洋的各個部分分配權限的大量管理開銷。」



架構典範對決：重新拿回您的資料自主權

維度	傳統架構 (Legacy SIEM)	解耦架構 (Cribl Decoupled)
資料攝取路徑	直接強制綁定單一節點	智慧路由，適配最佳目標
儲存成本	昂貴的 Premium 階層 (Pay-by-Ingest)	低成本分層儲存 (高達 85% 壓縮率)
供應商鎖定	高度綁定，轉移困難	完全中立 (Agnostic)，無縫切換平台
歷史資料調閱	重複攝取，耗時且昂貴	獨家 Replay 一鍵隨需回放

結論：Cribl 並非取代您的 SIEM，而是為您的 SIEM 裝上智慧引擎，實現「花費更少，看見更多」的終極目標。